

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 07:13:21 ; Search time 1916 Seconds

(without alignments)  
13078.035 Million cell updates/sec

Title: US-09-934-249-3

Sequence: 1 atgcaccgccttgatggtgggtc.....aacagaagaagcaccctctc 861

Scoring table: OLIGO-MNC

Gapop 60.0 , Capext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*  
29: em.vi:\*  
30: em.htg.hum:\*  
31: em.htg.in:\*  
32: em.htg.other:\*  
33: em.htg.mus:\*  
34: em.htg.pln:\*  
35: em.htg.rtd:\*  
36: em.htg.mam:\*  
37: em.htg.vrt:\*  
38: em.sy:\*  
39: em.htgo.hum:\*  
40: em.htgo.mus:\*  
41: em.htgo.other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861	100.0	861	6 AX392419	AX392419 Sequence
2	861	100.0	1321	6 AX392417	AX392417 Sequence
3	861	100.0	4839	9 AF305616	AF305616 Homo sapi
4	790	91.8	1061	9 BC015918	BC015918 Homo sapi
5	752	87.3	1141	9 AF224278	AF224278 Homo sapi
6	752	87.3	1818	9 AY128643	AY128643 Homo sapi
7	545	63.3	61505	9 AF305426	AF305426 Homo sapi
8	545	63.3	130435	9 HS71807	HS71807 Homo sapi
9	401	46.6	408	6 AX071267	AX071267 Sequence
10	302	35.1	693	6 AX392430	AX392430 Sequence
11	271	31.5	812	6 AX011709	AX011709 Sequence
12	156	18.1	150224	9 HSJ105917	HSJ105917 Sequence
13	51	5.9	51	6 AX199565	AX199565 Sequence
14	51	5.1	651	10 AF220208	AF220208 Mus muscu
15	44	5.1	878	6 AX392428	AX392428 Sequence
16	44	5.1	156075	2 AL837509	AL837509 Mus muscu
17	44	5.1	167758	2 AC110189	AC110189 Mus muscu
18	41	4.8	475	6 AX392431	AX392431 Sequence
19	34	3.9	159824	2 AC111878	AC111878 Rattus no
20	32	3.7	756	10 AB045588S2	AB045588 Rattus no
21	32	3.7	56475	9 AL353658	AL353658 Human DNA
22	32	3.7	60989	2 AC094865	AC094865 Rattus no
23	32	3.7	111961	2 AC126194	AC126194 Rattus no
24	32	3.7	163085	2 AC106160	AC106160 Rattus no
25	32	3.7	165438	2 AC126988	AC126988 Rattus no
26	32	3.7	172525	2 AC008514	AC008514 Homo sapi
27	32	3.7	179372	2 AC040922	AC040922 Homo sapi
28	32	3.7	198582	2 AC005291	AC005291 Homo sapi
29	32	3.7	223029	2 AC111410	AC111410 Rattus no
30	31	3.6	67987	10 AL672157	AL672157 Mouse DNA
31	31	3.6	70811	2 AC117542	AC117542 Mus muscu
32	31	3.6	96240	2 AC110445	AC110445 Rattus no
33	31	3.6	114287	2 AP000590	AP000590 Homo sapi
34	31	3.6	162812	2 AC129076	AC129076 Mus muscu
35	31	3.6	166474	2 AC107431	AC107431 Rattus no
36	31	3.6	169163	2 AL606807	AL606807 Homo sapi
37	31	3.6	183301	2 AC092487	AC092487 Homo sapi
38	31	3.6	189982	2 AC079023	AC079023 Mus muscu
39	31	3.6	210608	2 AC006028	AC006028 Homo sapi
40	30	3.5	79951	9 HSJ908M14	HSJ908M14 Human DNA
41	30	3.5	85546	2 AC111932	AC111932 Rattus no
42	30	3.5	99497	9 HS29316	HS29316 Human DNA
43	30	3.5	145241	2 AC117110	AC117110 Rattus no
44	30	3.5	156013	2 AC126882	AC126882 Rattus no
45	30	3.5	167936	2 AC106143	AC106143 Rattus no

#### ALIGNMENTS

RESULT 1	AX392419	861 bp	DNA	Linear	PAT 23-MAR-2002
LOCUS	AX392419				
DEFINITION	Sequence 3 from Patent WO0216416.				
ACCESSION	AX392419				
VERSION	AX392419.1	GI:19700734			
KEYWORDS					
SOURCE					
ORGANISM	human.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1				
AUTHORS	Iee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and				
TITLE	Turi,T.G.				
	Diagnosis and treatment of cardiovascular conditions				

Pred. No. is the number of results predicted by chance to have a





QY 841 AACAGAAAGACACCTCTC 861  
 Db 1161 AACAGAAAGACACCTCTC 1181  
 RESULT 4  
 LOCUS BC015918 1061 bp mRNA linear PRI 17-OCT-2001  
 DEFINITION Homo sapiens, clone MGC:20374 IMAGE:4559576, mRNA, complete cds.  
 ACCESSION BC015918  
 VERSION BC015918.1 GI:16198474  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1061)  
 Strausberg, R.  
 Direct Submission  
 Submitted (15-OCT-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: cgapsd-remail.nih.gov  
 Tissue Procurement: DCTD/DRP  
 CDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@cgsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven  
 Ness, Pawan Pandoh, Anna-Lissa Prabhu, Parvaneh Seedei, Jacqueline  
 Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,  
 Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Series: IRAL Plate: 29 Row: e Column: 5  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Similarity but not  
 identity to protein.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..1061  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:20374 IMAGE:4559576"  
 /rissue\_type="Kidney, renal cell adenocarcinoma"  
 /clone\_lib="NIH\_MGC\_14"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 131..844  
 /codon\_start=1  
 /product="unknown (protein for MGC:20374)"  
 /protein\_id="AAH15918.1"  
 /db\_xref="GI:16198475"  
 /translation="MMVAVVITCLSHYKLSARFISRHSGRRRDAASSGCLMP  
 SESTVSGNGIPEPOVYAPRPRTDLAVPRAQREPRRPPTPYIOHEIDLPPTLS  
 SDGEPPYOGCTIQRDPEOLELNREVRAPRNKTITDSULMSARLGGCPSS  
 NSGISATCVSGSGRMREPPPTTISEVIGHFPGSSFOHQSSGPPSLLEGRLHTHTIAP  
 LESATISKKRDKQKGPPL"  
 BASE COUNT 239 a 350 c 311 g 161 t  
 ORIGIN  
 Query Match 91.8%; Score 790; DB 9; Length 1061;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 840; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 CAACAGCACCAGCCGCGCCGCGGAGGCCAATGTCCTCTGACAGTGCATCGCA 80  
 Db 1 CAACAGCACCAGCCGCGCCGCGGAGGCCAATGTCCTCTGACAGTGCATCGCA 60  
 QY 81 ACCGCTCTTTGTCACAGCATGAGATCAACGAGCTGGAGTTGTTCAAGATCATCAT 140  
 Db 61 ACCGCTCTTTGTCACAGCATGAGATCAACGAGCTGGAGTTGTTCAAGATCATCAT 120  
 QY 141 CGTGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 200  
 Db 121 CGTGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 QY 201 GTCTGCACGCTCTTCATCAGCCGACAGCCAGGAGGAGAGAGAGAGAGAGAGAGAG 260  
 Db 181 GTCTGCACGCTCTTCATCAGCCGACAGCCAGGAGGAGAGAGAGAGAGAGAGAGAG 240  
 QY 261 CTCAGAGAGATGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 320  
 Db 241 CTCAGAGAGATGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 321 GCAGGTCTAGCCGCGCCGCTGCGCCACAGCAGCCGCTGCGCCGCTGCGCCAGCG 380  
 Db 301 GCAGGTCTAGCCGCGCCGCTGCGCCACAGCAGCCGCTGCGCCGCTGCGCCAGCG 360  
 QY 381 GGAGGCTTCCACCGCTTCCAGCCACCTATCCGTAAGCAGCAGATGACCTGCGC 440  
 Db 361 GGAGGCTTCCACCGCTTCCAGCCACCTATCCGTAAGCAGCAGATGACCTGCGC 420  
 QY 441 ACCCAGCATCTGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 500  
 Db 421 GCCCAGCATCTGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 QY 501 CCAGCTTGGGAGCCCGAGCAGCAGCTGGAATGAACCGGAGTGGTGGCGCAGCCCG 560  
 Db 481 CCAGCTTGGGAGCCCGAGCAGCAGCTGGAATGAACCGGAGTGGTGGCGCAGCCCG 540  
 QY 561 AAACGAGACATCTTGCAGCACTGACAGCTGATGATGATGATGATGATGATGATGAT 620  
 Db 541 AAACGAGACATCTTGCAGCACTGATGATGATGATGATGATGATGATGATGATGAT 600  
 QY 621 CCCGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680  
 Db 601 CCCGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 QY 681 GCGCGCGCCACCTACAGCGAGGTGATGCGCACTACCGGGGTCTCTTCCAGCAGCA 740  
 Db 661 GCGCGCGCCACCTACAGCGAGGTGATGCGCACTACCGGGGTCTCTTCCAGCAGCA 720  
 QY 741 GCAGAGCATGTGGCGCGCTCTCTGCTGAGAGGAGACCGGCTCCACACACACATGCC 800  
 Db 721 GCAGAGCATGTGGCGCGCTCTCTGCTGAGAGGAGACCGGCTCCACACACATGCC 780  
 QY 801 GCCCTTGAAGAGCGCAGCATCTGTGAGCAAGAGAGATTAACAGAAAGACACCTCT 860  
 Db 781 GCCCTTGAAGAGCGCAGCATCTGTGAGCAAGAGAGATTAACAGAAAGACACCTCT 840  
 QY 861 C 861  
 Db 841 C 841  
 RESULT 5  
 AF224278 1141 bp mRNA linear PRI 18-JUL-2000  
 LOCUS Homo sapiens PMEPAL protein (PMEPAL) mRNA, complete cds.  
 DEFINITION AF224278  
 ACCESSION AF224278.1 GI:9255808  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.



```

REFERENCE 1 (bases 1 to 1141)
AUTHORS Xu,L.L., Shanmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,
TITLE A novel androgen-regulated gene, PMEPA1, located on chromosome
JOURNAL 20q13 exhibits high level expression in prostate
MEDLINE 20334621
PUBMED 10873380
REFERENCE 2 (bases 1 to 1141)
AUTHORS Xu,L.L., Shanmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,
TITLE A novel androgen-regulated gene, PMEPA1, located on chromosome
JOURNAL 20q13 exhibits high level expression in prostate
MEDLINE 20334621
PUBMED 10873380
FEATURES
    source             location/Qualifiers
                        1..1141
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="20"
                        /map="20q13.31-13.33"
                        /cell_line="LNCaP"
                        1..1141
                        /gene="PMEPA1"
                        /note="type 1b transmembrane protein; expression is
                        induced in response to the synthetic androgen, R1881;
                        expression is abundant in, and restricted to, prostate
                        glandular epithelial cells; similar to the predicted
                        protein encoded by sequence deposited at GenBank Accession
                        Number NP_004329"
                        /codon_start=1
                        /product="PMEPA1 protein"
                        /protein_id="AA86322.1"
                        /db_xref="GI:9255809"
                        /translation="MALEFVQIIIVVVMVVVITCLSHYKLSARSFISRHSG
                        RREDLSSBGLPESVSGNIPPEQVYAPRPDRLAVPEPQREHFRPQY
                        PYLQHEIDLPPTISLDGEEPPYQPCITOLDPEQLEINRESVAPNRTIFSD
                        LMSARLGGPCPPSSNSGTSATCYSGSGRREGPPPTSEVIGHYRGSSFOHOSSGPP
                        SLLEGRRLHHTHAPLESAAIWSKEXDKQGHPL"
BASE COUNT      271 a      350 c      336 g      184 t
ORIGIN
Query Match      87.3%; Score 752; DB 9; Length 1141;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 110 CGAGCTGAGAGTTGTCAGATCATCATCTGCTGATGATGATGATGATG 169
    |||||||
DB 100 CGAGCTGAGAGTTGTCAGATCATCATCTGCTGATGATGATGATGATG 159
QY 170 TGATCAGTGCCTGCTGAGCAGTACAAAGCTGTCTGACAGGCTCTCATCAGCCGGCACA 229
    |||||||
DB 160 TGATCAGTGCCTGCTGAGCAGTACAAAGCTGTCTGACAGGCTCTCATCAGCCGGCACA 219
QY 230 GCCAGGGGGCGAGAGAGAATATCCCTCTCTCAGAAAGATGCTGTGGCCCTTGGAGA 289
    |||||||
DB 220 GCCAGGGGGCGAGAGAGAATATCCCTCTCTCAGAAAGATGCTGTGGCCCTTGGAGA 279
QY 290 GCACAGTGTCAAGGACAGGATCCCAAGAGCGGAGGTGTACGCCCGGCTGGGCCACAG 349
    |||||||
DB 280 GCACAGTGTCAAGGACAGGATCCCAAGAGCGGAGGTGTACGCCCGGCTGGGCCACAG 339
QY 350 ACCGCTGGCCGTCGCCCTTCCGCCAGCGGAGCGGCTTCCAGCGCTTCCAGCCACCT 409
    |||||||
DB 340 ACCGCTGGCCGTCGCCCTTCCGCCAGCGGAGCGGCTTCCAGCGCTTCCAGCCACCT 399
QY 410 ATCCGTACTCTGACAGCAGATCGACCTGCCACCCACCATCTCGTGTCAACAAGGGGAGG 469
    |||||||
DB 400 ATCCGTACTCTGACAGCAGATCGACCTGCCACCCACCATCTCGTGTCAACAAGGGGAGG 459
QY 470 AGCCCCACCTTACAGAGGGCCCTGCACCTTCAGAGCTTGGGAGCCCGAGCAGCAGCTGG 529
    |||||||

```

```

DB 460 AGCCCCACCTTACAGAGGGCCCTGTGACACCTTCAGCTTCGGGAGCCCGAGCAGCAGCTGG 519
QY 530 AACTGAACCGGAGTCTGCTGGCGGACCGCCCAACAGACCATCTTGACAGTACCTGA 589
    |||||||
DB 520 AACTGAACCGGAGTCTGCTGGCGGACCGCCCAACAGACCATCTTGACAGTACCTGA 579
QY 590 TGCATAGTGCAGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 649
    |||||||
DB 580 TGCATAGTGCAGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 639
QY 650 CTGCTACAGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 709
    |||||||
DB 640 CTGCTACAGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 699
QY 710 GCCACTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 769
    |||||||
DB 700 GCCACTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 759
QY 770 AGGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 829
    |||||||
DB 760 AGGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 819
QY 830 AAGAGAGATTAACAGAAAGACACCTCTC 861
    |||||||
DB 820 AAGAGAGATTAACAGAAAGACACCTCTC 851
RESULT 6
AY128643      1818 bp      mRNA      linear      PRI 05-AUG-2002
LOCUS        Homo sapiens PMEPA1 variant A protein mRNA, complete cds.
DEFINITION   AY128643
ACCESSION    AY128643
VERSION      AY128643.1 GI:22121998
KEYWORDS     human.
ORGANISM     Homo sapiens
              Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 1818)
              Brunschwig,E.B., Wilson,K., Mack,D., Dawson,D., Lawrence,E.,
              Willison,J.K.V., Lu,S., Nosrati,A., Swinler,S., Beard,L.,
              Lutterbaugh,J.D., Willis,J., Platzer,P. and Markowitz,S.
              Direct Submission
              Submitted (03-JUL-2002) Department of Medicine, Case Western
              Reserve University/Howard Hughes Medical Institute, 11001 Cedar
              Ave., Cleveland, OH 44106, USA
TITLE        Submitted (03-JUL-2002) Department of Medicine, Case Western
JOURNML      Reserve University/Howard Hughes Medical Institute, 11001 Cedar
              Ave., Cleveland, OH 44106, USA
FEATURES
    source             location/Qualifiers
                        1..1818
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="20"
                        /map="20q13.31-13.33"
                        145..858
                        /codon_start=1
                        /product="PMEPA1 variant A protein"
                        /protein_id="AA86927.1"
                        /db_xref="GI:22121998"
                        /translation="MWVVVVVITCLSHYKLSARSFISRHSGRREDLSSBGLP
                        SESVSGNIPPEQVYAPRPDRLAVPEPQREHFRPQY
                        PYLQHEIDLPPTISLDGEEPPYQPCITOLDPEQLEINRESVAPNRTIFSDLMSARLGGPCPP
                        SSNSGTSATCYSGSGRREGPPPTSEVIGHYRGSSFOHOSSGPPSLBGRRLHHTHAP
                        LESAAIWSKEXDKQGHPL"
BASE COUNT      459 a      467 c      494 g      398 t
ORIGIN
Query Match      87.3%; Score 752; DB 9; Length 1818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 110 CGAGCTGAGAGTTGTCAGATCATCATCTGCTGATGATGATGATGATG 169
    |||||||
DB 104 CGAGCTGAGAGTTGTCAGATCATCATCTGCTGATGATGATGATGATG 163

```

[illegible]

```

source
1. .61505
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.2-q13.33"
1. .61505
/gene="STAG1/PMEPA1"
join1. .429,50206. .50360,56817. .56870,57305. .61505)
/gene="STAG1/PMEPA1"
/product="solid tumor-associated 1 protein"
join1321. .429,50206. .50360,56817. .56870,57305. .57850)
/gene="STAG1/PMEPA1"
/codon_start=1
/product="solid tumor-associated 1 protein"
/protein_id="AAL09357.1"
/db_xref="GI:15824469"
/translation="MHRIMGVNSTAAAGAOPNVSCGCKNRSLFQSMETELFEVQ
IIIVVAVMVAIVITCLLSHKLSARSTISRHSGQRREDALSSGCLMPSSSTVSG
GIPEVQVAVPRPDRILAVPFAQERERERQPTVPLQHEIDLPTTSLSGEPEP
YQGGTGLRDPEOLELNRSEVRAPRPNRTIPDSLDMSARLGCGPPSSNGISATQ
YSGSGRMGEGPPTYSVETIGHYFGSSFGHQSSGPPSLLEGITLHTHTIAPLESAIIM
KEKKOKKHP"
BASE COUNT      13329 a      16074 c      17430 g      14672 t
ORIGIN

```

Query Match	63.3%	Score 545	DB 9	Length 61505
Best Local Similarity	100.0%	Pred. NO. 4.2e-275		
Matches 545	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	317	AGCCGAGGTCTACGGCCCCGCTGGGCCACAGACCGCCCTGGGCGGTGGCCCTTCGCCC	376	
Db	57303	AGCCGAGGTCTACGGCCCCGCTGGGCCACAGACCGCCCTGGGCGGTGGCCCTTCGCCC	57362	
QY	377	AGCGGAGGGCTTCCACCCGCTTCAGACCACCTATCCGTAAGTCTGCAGCAGAGATGACC	436	
Db	57363	AGCGGAGGGCTTCCACCCGCTTCAGACCACCTATCCGTAAGTCTGCAGCAGAGATGACC	57422	
QY	437	TGCGAACCCACATCTGCTGTGTCAGAGGGGAGAGAGGCCCCACCTTACAGGGGCCCTTGA	496	
Db	57423	TGCGAACCCACATCTGCTGTGTCAGAGGGGAGAGAGGCCCCACCTTACAGGGGCCCTTGA	57482	
QY	497	CCCTCAGCTTGTGGGACCCCGAGCAGACGTGAACTGAACCGGGAGTGGTGGCGCAC	556	
Db	57483	CCCTCAGCTTGTGGGACCCCGAGCAGACGTGAACTGAACCGGGAGTGGTGGCGCAC	57542	
QY	557	CCCGAAACGAACCATCTTCGACAGGACCTGATGGATGTGTCAGGGTGGGCGGCCCT	616	
Db	57543	CCCGAAACGAACCATCTTCGACAGGACCTGATGGATGTGTCAGGGTGGGCGGCCCT	57602	
QY	617	GGCCCCCGCAGCTACTCGGGCATCAGCGCCACGCTACGGCAGCGCGGGCGCATGG	676	
Db	57603	GGCCCCCGCAGCTACTCGGGCATCAGCGCCACGCTACGGCAGCGCGGGCGCATGG	57662	
QY	677	AGGGCGCGCGCCCACTACAGCAGAGTCATGGGCCACTACCGGGGGTCTCTTCCAGC	736	
Db	57663	AGGGCGCGCGCCCACTACAGCAGAGTCATGGGCCACTACCGGGGGTCTCTTCCAGC	57722	
QY	737	ACCACGAGGCGATGGGGCGCCCTCTCTGTCGTGAGAGGGACCGGCTCCACACACACACA	796	
Db	57723	ACCACGAGGCGATGGGGCGCCCTCTCTGTCGTGAGAGGGACCGGCTCCACACACACACA	57782	
QY	797	TGCGCGCCCTAGAGAGCGCAGCATCTGGAGCAAGAGAGATTAACAGAAAGACACC	856	
Db	57783	TGCGCGCCCTAGAGAGCGCAGCATCTGGAGCAAGAGAGATTAACAGAAAGACACC	57842	
QY	857	CTCTC 861		
Db	57843	CTCTC 57847		
RESULT 8				
HS718J7/c				
LOCUS	HS718J7	130435 bp	DNA	linear
				PRI 24-FEB-2001

DEFINITION Human DNA sequence from clone RP4-718J7 on chromosome 20q13.31-13.33 contains the PCXI gene, part of a novel phosphoenolpyruvate carboxykinase 1, part of a novel gene similar to mouse DLM-1 (tumour stroma and activated macrophage protein), the 3' end of the TMEPAI gene encoding an androgen induced 1b transmembrane protein (PMEPAI), two putative novel genes, a Cpg Island, ESTs, STS and GSSs, complete sequence.

ACCESSION AL035541

VERSION AL035541.15 GI:11546043

KEYWORDS HTG; Cpg Island; DLM-1; macrophage protein; PCXI; phosphoenolpyruvate carboxykinase; PMEPAI; TMEPAI; transmembrane protein.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 130435)

AUTHORS Sehra,H.

JOURNAL Direct Submission

COMMENT Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Dec 5, 2000 this sequence version replaced gi:10198628. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/c-elegans/wormpep>

This sequence is the entire insert of clone RP4-718J7 The true left end of clone RP5-1007E6 is at 71437 in this sequence. The true right end of clone RP4-579F20 is at 43945 in this sequence. RP4-718J7 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES

source

1..130435

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="20"

/map="q13.31-13.33"

/clone="RP4-718J7"

/clone\_1lb="RPCI-4"

repeat\_region

319..478

/note="5 copies 32 mer 67% conserved"

repeat\_region

370..463

/note="2 copies 47 mer 87% conserved"

repeat\_region

427..626

/note="4 copies 50 mer 71% conserved"

repeat\_region

707..812

/note="53 copies 2 mer cc 61% conserved"

repeat\_region

967..2752

/note="893 copies 2 mer gg 54% conserved"

repeat\_region

982..2730

/note="33 copies 53 mer 54% conserved"

repeat\_region

1177..1820

/note="4 copies 161 mer 64% conserved"

repeat\_region

1204..1371

/note="3 copies 56 mer 75% conserved"

repeat\_region

1358..2757

/note="28 copies 50 mer 54% conserved"

repeat\_region

1675..2718

/note="18 copies 58 mer 55% conserved"

repeat\_region

1843..2346

/note="9 copies 56 mer 64% conserved"

repeat\_region

1852..2079

/note="4 copies 57 mer 86% conserved"

repeat\_region

2028..2632

/note="11 copies 55 mer 60% conserved"

repeat\_region

2087..2257

/note="3 copies 57 mer 83% conserved"

repeat\_region

2228..2497

/note="5 copies 54 mer 75% conserved"

repeat\_region

2583..2716

/note="2 copies 67 mer 82% conserved"

repeat\_region

3378..3426

/note="LIMB4 repeat: matches 6088..6136 of consensus"

repeat\_region

3799..3896

/note="Charliel repeat: matches 681..781 of consensus"

repeat\_region

5331..5793

/note="MTR1B repeat: matches 14..466 of consensus"

repeat\_region

5797..5988

/note="6 copies 32 mer 86% conserved"

repeat\_region

5855..5978

/note="31 copies 4 mer gcac 61% conserved"

repeat\_region

6474..6591

/note="MIR repeat: matches 91..218 of consensus"

repeat\_region

6592..6723

/note="FLAK\_C repeat: matches 1..132 of consensus"

repeat\_region

6952..7021

/note="L2 repeat: matches 2637..2705 of consensus"

repeat\_region

7358..7671

/note="AluX repeat: matches 1..312 of consensus"

repeat\_region

8521..8554

/note="Alu repeat: matches 1..34 of consensus"

repeat\_region

8863..9217

/note="MTR1D repeat: matches 105..505 of consensus"

repeat\_region

9228..9746

/note="L2 repeat: matches 1459..2028 of consensus"

repeat\_region

9880..10089

/note="L2 repeat: matches 2292..2511 of consensus"

repeat\_region

10094..10206

/note="AluB repeat: matches 188..300 of consensus"

repeat\_region

10263..10573

/note="AluB repeat: matches 1..308 of consensus"

repeat\_region

10658..10753

/note="L2 repeat: matches 2647..2749 of consensus"

repeat\_region

11816..12380

/note="LRR19B repeat: matches 1..580 of consensus"

repeat\_region

12519..12813

/note="LIMD2 repeat: matches 6032..6331 of consensus"

repeat\_region

13559..13726

/note="3 copies 56 mer 73% conserved"

repeat\_region

13585..13716

/note="22 copies 6 mer cttct 67% conserved"

repeat\_region

13586..13717

/note="66 copies 2 mer tc 68% conserved"

repeat\_region

13588..13715

/note="4 copies 32 mer 78% conserved"

repeat\_region

13589..13724

/note="34 copies 4 mer cttt 77% conserved"

misc-feature

complement(13622..14142)

/note="match: GSS: Em:AQ592603"

repeat\_region

14313..14437

/note="LRR16C repeat: matches 257..386 of consensus"

repeat\_region

15015..15110

/note="3 copies 32 mer 79% conserved"

repeat\_region

15021..15110

/note="45 copies 2 mer ac 80% conserved"

```
repeat_region 15026..15109
/note="7 copies 12 mer 82% conserved"
repeat_region 15028..15111
/note="14 copies 6 mer cacaca 82% conserved"
repeat_region 15029..15108
/note="20 copies 4 mer acac 83% conserved"
repeat_region 15273..15399
/note="L2 repeat: matches 2159..2285 of consensus"
repeat_region 15420..15721
/note="AluX repeat: matches 1..303 of consensus"
repeat_region 16333..16396
/note="2 copies 32 mer 98% conserved"
repeat_region 16575..16690
/note="L2 repeat: matches 2596..2711 of consensus"
repeat_region 18049..18169
/note="MIR repeat: matches 86..211 of consensus"
repeat_region 18312..18438
/note="MIR repeat: matches 138..250 of consensus"
repeat_region 18585..18776
/note="MER20 repeat: matches 7..217 of consensus"
repeat_region 19211..19290
/note="MIR repeat: matches 72..154 of consensus"
repeat_region 20413..20463
/note="MIR repeat: matches 90..143 of consensus"
repeat_region 21672..21672
/note="match: GSS: Em: A0748384"
repeat_region 21672..21672
/note="match: GSS: Em: A076209"
repeat_region 21667..21775
/note="match: GSS: Em: A0892114"
repeat_region 21667..21775
/note="MIR repeat: matches 73..192 of consensus"
repeat_region 23577..23646
/note="2 copies 35 mer 100% conserved"
repeat_region 24257..24591
/note="L2 repeat: matches 2336..2710 of consensus"
repeat_region 25580..25611
/note="16 copies 2 mer tc 87% conserved"
repeat_region 26333..26643
/note="Alu repeat: matches 1..310 of consensus"
repeat_region 27603..27684
/note="L2 repeat: matches 2661..2739 of consensus"
repeat_region 27823..28041
/note="MIR repeat: matches 38..242 of consensus"
repeat_region 28129..28259
/note="FLAN_C repeat: matches 1..127 of consensus"
repeat_region 31203..31356
/note="LMB5 repeat: matches 6015..6176 of consensus"
repeat_region 32224..32522
/note="AluX repeat: matches 1..298 of consensus"
repeat_region 33285..33481

Query Match 63.3%; Score 545; DB 9; Length 130435;
Best Local Similarity 100.0%; Pred. No. 4.2e-275;
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 557 CCCCAAGAACACATCTTCAGACGTGACGACCTGATAGTCGACGGTGGGCGCCCT 616
|
|
|
Db 128203 CCCCAAGAACACATCTTCAGACGTGACGACCTGATAGTCGACGGTGGGCGCCCT 128144
|
|
|
QY 617 GCCCCCCAGCAGTAATCTGGGATCAGGGCCACGCTGTACGGCAGGGGGCGCATGG 676
|
|
|
Db 128143 GCCCCCCAGCAGTAATCTGGGATCAGGGCCACGCTGTACGGCAGGGGGCGCATGG 128084
|
|
|
QY 677 AGGGGCGCGCGCCACCTACAGGAGTTCATGGGCCTACCCGGGGTCTCTCTCCAGC 736
|
|
|
Db 128083 AGGGGCGCGCGCCACCTACAGGAGTTCATGGGCCTACCCGGGGTCTCTCTCCAGC 128024
|
|
|
QY 737 ACCACAGACAGCTGGGCGCGCCCTCTGCTGGAGGGAGCCGGCTCCACACACACA 796
|
|
|
Db 128023 ACCACAGACAGCTGGGCGCGCCCTCTGCTGGAGGGAGCCGGCTCCACACACACA 127964
|
|
|
QY 797 TCGCGCCCTAGAGAGCGGACCATCTGGAGCAAGAGATTAACAGAAAGACAC 856
|
|
|
Db 127963 TCGCGCCCTAGAGAGCGGACCATCTGGAGCAAGAGATTAACAGAAAGACAC 127904
|
|
|
QY 857 CTCTC 861
|
|
|
Db 127903 CTCTC 127899
|
|
|

RESULT 9
AX071267 408 bp DNA linear PAT 25-JAN-2001
LOCUS Sequence 1739 from Patent WO0102568.
DEFINITION AX071267
ACCESSION AX071267
VERSION AX071267.1 GI:12581618
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 408)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
Kassam,A., Reinhard,C., Randoz,F., Kennedy,G.C., Pet,D.,
Lamson,G., Dimaio,R., Gikenjakov,R., Dimaio,S., Dickson,M.,
Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
Human genes and gene expression products
Patent: WO 0102568-A 1739 11-JAN-2001;
JOURNAL CHIRON CORPORATION (US); HYSQ, INC. (US)
FEATURES
location/qualifiers
1..408
SOURCE
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 81 a 159 c 115 g 53 t
ORIGIN

Query Match 46.6%; Score 401; DB 6; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.3e-199;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Db 248 TCCGCTACTGACGACGAGATGACAGCTGCGACCCACCATCTCGGTGTACAGCGGGAGGA 307
QY 471 GCGCCCACTACGACGAGGCGCCCTGCACCTCTGACCTGGAGACCCGACGACGCTGGA 530
Db 308 GCGCCCACTACGACGAGGCGCCCTGCACCTCTGACCTGGAGACCCGACGAGCTGGA 367
QY 531 ACTGACCGGAGTGTGCTGCGCGACCCCAACAGAACCA 571
Db 368 ACTGACCGGAGTGTGCTGCGCGACCCCAACAGAACCA 408

RESULT 10
AX392430/c 693 bp DNA linear PAT 23-MAR-2002
LOCUS Sequence 14 from Patent W00216416.
ACCESSION AX392430
VERSION AX392430.1 GI:19700746
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Lee,R.T., Landschultz,K.T., Kennedy,S.P., Thompson,J.F. and
Turli,T.G.
TITLE Diagnosis and treatment of cardiovascular conditions
JOURNAL Patent: WO 0216416-A 14 28-FEB-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
FEATURES
source 1..693
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 101 a 205 c 237 g 149 t 1 others
ORIGIN

Query Match 35.1%; Score 302; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 2.8e-147;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 CCCCACCTTACGAGGCGCCCTGCACCTTCCAGCTTGGGAGCCCGAGCAGCAGCTGGAC 532
Db 587 CCCCACCTTACGAGGCGCCCTGCACCTTCCAGCTTGGGAGCCCGAGCAGCAGCTGGAC 528
QY 533 TGAACCGGAGTGTGCTGCGCGACCCCAACAGAACCATTTGAGACGTGACCTGATGG 592
Db 527 TGAACCGGAGTGTGCTGCGCGACCCCAACAGAACCATTTGAGACGTGACCTGATGG 468
QY 593 ATAGTCCAGGCTGGGCGGCGCCCTGCACCTTAACTGGGCTACAGCGCCACAGT 652
Db 467 ATAGTCCAGGCTGGGCGGCGCCCTGCACCTTAACTGGGCTACAGCGCCACAGT 408
QY 653 GCTACGCGACGCGGCGGCGCATGTAGGAGGCGCGCCCACTACAGCAGAGGTATCGGCC 712
Db 407 GCTACGCGACGCGGCGGCGCATGTAGGAGGCGCGCCCACTACAGCAGAGGTATCGGCC 348
QY 713 ACTACCGGCGGCTCTCTTCCAGACGACGAGCAGAGCGGCGCCCTCTCTTGTGTGAGG 772
Db 347 ACTACCGGCGGCTCTCTTCCAGACGACGAGCAGAGGTGCGCGCCCTCTTGTGTGAGG 288
QY 773 GG 774
Db 287 GG 286

RESULT 11
AX011709 812 bp DNA linear PAT 06-SEP-2000
LOCUS Sequence 107 from Patent W09955858.
ACCESSION AX011709
VERSION AX011709.1 GI:9998233
KEYWORDS
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 812)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarzky,C.
TITLE Human nucleic acid sequences obtained from pancreas tumor tissue
JOURNAL Patent: WO 9955858-A 107 04-NOV-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METZGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source 1..812
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 157 a 272 c 237 g 146 t
ORIGIN

Query Match 31.5%; Score 271; DB 6; Length 812;
Best Local Similarity 99.7%; Pred. No. 5.6e-131;
Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 173 TCACGTGCTGCTGAGCCACTACAGCTGTCTGACAGGCTCTTCATACGCCGACAGCC 232
Db 426 TCACGTGCTGCTGAGCCACTACAGCTGTCTGACAGGCTCTTCATACGCCGACAGCC 485
QY 223 AGGGCGGAGAGAGAGATGCTGCTCTCAGAAAGATGCTGTGGCCCTCGGAGAGCA 292
Db 486 AGGGCGGAGAGAGAGATGCTGCTCTCAGAAAGATGCTGTGGCCCTCGGAGAGCA 545
QY 293 CAGTGTGAGGACAGGAATCCAGAGCGGAGGCTACAGCCCGGCTGCGGCGACCGAGC 352
Db 546 CAGTGTGAGGACAGGAATCCAGAGCGGAGGCTACAGCCCGGCTGCGGCGACCGAGC 605
QY 353 GCTTGGCGGCTGCGGCTTCCGACGCGGAGGCGCTTCCAGCCGACCTATATC 412
Db 606 GCTTGGCGGCTGCGGCTTCCGACGCGGAGGCGCTTCCAGCCGACCTATATC 665
QY 413 CGTACCTGACGACGAGATGACCTGCACCCACCATCTGCTGTACAGCGGGAGAGC 472
Db 666 CGTACCTGACGACGAGATGACCTGCACCCACCATCTGCTGTACAGCGGGAGAGC 725
QY 473 CCCCACCTTACGAGGCGCCCTGC 494
Db 726 CCCCACCTTACGAGGCGCCCTG 747

RESULT 12
HSJ10591L/c 150224 bp DNA linear PRI 24-FEB-2001
LOCUS Human DNA sequence from clone R95-10591L on chromosome
20q13.2-13.33 contains the 5' end of the TMEPA1 (PMEPA1) gene
encoding an androgen induced 1b transmembrane protein, ESTs, STSs,
GSSs and two CpG islands, complete sequence.
ACCESSION AL121913
VERSION AL121913.4 GI:7161781
KEYWORDS HMG; CpG island; PMEPA1; TMEPA1; transmembrane protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 150224)
AUTHORS Skuce,C.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Mar 6, 2000 this sequence version replaced gi:7007305.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with

```

only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/RGP/Chr20>

This sequence is the entire insert of clone RP5-1059L7. The true left end of clone RP11-402F1 is at 106677 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-1059L7 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: PCYPAC2

## FEATURES

### source

Location/Qualifiers

1..150224

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="20"

/map="q13.2-13.33"

/clone="RP5-1059L7"

/clone\_id="RPCI-5"

283..355

/note="MIR repeat: matches 71..145 of consensus"

1509..1628

/note="MIR repeat: matches 46..165 of consensus"

2076..2155

/note="L2 repeat: matches 2616..2707 of consensus"

complement(2445..53425)

/gene="TMEPA1"

complement(join(<2445..2599,53362..53425))

/gene="TMEPA1"

/product="dJ1059L7.1.2 (androgen induced type 1b

transmembrane protein (PMEPA1), isoform 2)"

/note="match: CDNs: Em:AF224278

match: ESTs: Em:AA088767"

/evidence="not\_experimental"

complement(join(<2445..2599,53362..53365))

/gene="TMEPA1"

/note="continues in dJ718J7 (AL035541)"

/codon\_start=1

/evidence="not\_experimental"

/product="dJ1059L7.1.1 (androgen induced type 1b

transmembrane protein (PMEPA1), isoform 1)"

/protein\_id="CAC32857.1"

/db\_xref="GI:13160408"

/translation="VALEFVQIIIVVMVMVITCLSHKLSARSRISRHSG

RRREDALSS"

complement(join(<2445..2599,53376..52477))

/gene="TMEPA1"

/product="dJ1059L7.1.1 (androgen induced type 1b

transmembrane protein (PMEPA1), isoform 1)"

/note="match: CDNs: Em:AF009426 Em:AF009427 Em:AF009425

Em:AF009424

match: ESTs: Em:AA249792 Em:AI594390 Em:AA128075

Em:AA088767"

/evidence="not\_experimental"

complement(join(<2445..2599,52376..52472))

/gene="TMEPA1"

/note="continues in dJ718J7 (AL035541)"

match: proteins: Tr:O15166 Tr:O15168"

/codon\_start=1

/evidence="not\_experimental"

/product="dJ1059L7.1.1 (androgen induced type 1b

transmembrane protein (PMEPA1), isoform 1)"

/protein\_id="CAB88144.1"

/db\_xref="GI:7619746"

/db\_xref="SPTREMBL:O9NTR9"

/translation="KGVNSTAAAGAPVNSCTCNCKRSLFQSMETTELDFVQIIIV

VVMVMVITCLSHKLSARSRISRHSGRRREDALSS"

2538..2567

/note="10 copies 3 mer cac 90% conserved"

3128..3606

/note="match: GSS: Em:AQ059507"

3214..3324

/note="3 copies 37 mer 83% conserved"

3339..3413

/note="3 copies 25 mer 92% conserved"

3390..3921

/note="19 copies 28 mer 61% conserved"

3394..3893

/note="25 copies 20 mer 59% conserved"

3407..3892

/note="27 copies 18 mer 58% conserved"

3421..3910

/note="7 copies 70 mer 62% conserved"

3473..3904

/note="12 copies 36 mer 58% conserved"

3553..3932

/note="10 copies 38 mer 64% conserved"

3573..3697

/note="5 copies 25 mer 76% conserved"

3746..3921

/note="8 copies 22 mer 64% conserved"

4032..4069

/note="19 copies 2 mer 11% conserved"

4792..5001

/note="3 copies 70 mer 82% conserved"

5002..5190

/note="MER20 repeat: matches 1..200 of consensus"

5224..5282

/note="MIR repeat: matches 94..154 of consensus"

8103..8205

/note="MIR repeat: matches 48..161 of consensus"

complement(8858..9277)

/gene="TMEPA1"

/note="match: GSS: Em:AA0881699"

9224..9382

/note="MIR repeat: matches 81..239 of consensus"

9286..9706

/note="match: GSS: Em:AQ565377"

9295..9861

/note="match: GSS: Em:AQ757992"

10035..10140

/note="Charlied repeat: matches 39..146 of consensus"

11175..11343

/note="MIR repeat: matches 11..196 of consensus"

11630..12037

/note="L2 repeat: matches 2290..2705 of consensus"

12054..12362

/note="Alusx repeat: matches 1..311 of consensus"

14448..14493

/note="23 copies 2 mer at 76% conserved"

14634..14813

/note="MER20 repeat: matches 1..187 of consensus"

15096..15497

/note="Cpc island"

/evidence="not\_experimental"

15690..15935

/note="MIR repeat: matches 7..254 of consensus"

16989..17387

/note="MLT1A2 repeat: matches 1..374 of consensus"

17647..17703

/note="L2 repeat: matches 2679..2736 of consensus"

17699..18179

/note="match: GSS: Em:AQ703107"

17902..18039

ORIGIN		Query Match		5.9%	Score 51	DB 6	Length 51
Best Local Similarity		100.0%		Pred. No. 3e-15			
Matches 51		Conservative 0		Mismatches 0	Indels 0	Gaps 0	
QY	206	CACGGTCCTTCATATACGCGCGCACAGCCAGGGGCGGAGAGAGAGATGCC	256				
Db	1	CACGGTCCTTCATATACGCGCGCACAGCCAGGGGCGGAGAGAGATGCC	51				
<p>RESULT 14</p> <p>LOCUS AF220208</p> <p>DEFINITION Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.</p> <p>ACCESSION AF220208</p> <p>VERSION AF220208.1</p> <p>KEYWORDS GI:12004973</p> <p>SOURCE Mus musculus.</p> <p>ORGANISM Mus musculus</p> <p>REFERENCE 1 (bases 1 to 651)</p> <p>AUTHORS Jolliffe,C.N., Harvey,K.F., Haines,B.P., Parasivam,G. and Kumar,S.</p> <p>TITLE Identification of multiple proteins expressed in murine embryos as binding partners for the WW domains of the ubiquitin-protein ligase Nedd4</p> <p>JOURNAL Biochem. J. 351 Pt 3, 557-565 (2000)</p> <p>MEDLINE 20498735</p> <p>PUBMED 11042109</p>							

BASE COUNT	146 a	234 c	176 g	95 t	
ORIGIN					
Query Match	5.1%; Score 44; DB 10; Length 651;				
Best Local Similarity	100.0%; Pred. No. 1,5e-11;				
Matches	44; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
OY	463	GGGAGAGAGCCGCCACCTACAGGGCCCTGCACCCCTCCAGCT	506		
Db	223	GGGAGAGAGCCGCCACCTACAGGGCCCTGCACCCCTCCAGCT	266		
RESULT 15					
AX392428					
LOCUS	AX392428	878 bp	DNA	linear	PAT 23-MAR-2002
DEFINITION	Sequence 12 from Patent WO0216416.				
ACCESSION	AX392428				
VERSION	AX392428.1	GI:19700744			
KEYWORDS	.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

REFERENCE 1 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
AUTHORS Lee, R.T., Landschulz, K.T., Kennedy, S.P., Thompson, J.F. and  
Tui, T.G.  
TITLE Diagnosis and treatment of cardiovascular conditions  
JOURNAL Patent: WO 0216416-A 12 28-FEB-2002;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)  
FEATURES  
source 1. 878  
location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
20. .844  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD29011.1"  
/db\_xref="GI:19700745"  
/translation="MGVNGTAATAAAGQPNVSCACNCQRSLEPPSMETLEFVQIVIV  
VVMVVMVITCLSHYKLSARSPISRHSQARRDDGLSSEGLPSESTVSGMPEP  
QVAPRPRTDLAVPPFIQSRFOPTVYLOHEIALPPTISLSDGEPPPYGGCTLO  
LRDPEOLELNRESYRAPPNRTIFDSDLIDSTMIGGPGCPSSNSGISATCYSSGGME  
GPPPTYSVIGHYESSFQHOQNSPSSLEGTRLHSHIAPLENKEKOKGHP.L"  
BASE COUNT 179 a 305 c 247 g 147 t  
ORIGIN  
Query Match 5.1%; Score 44; DB 6; Length 878;  
Best Local Similarity 100.0%; Pred. No. 1.5e-11;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 463 GGGGAGAGCCCCCACCCTACCGAGGCCCCCTGCACCTCCAGCT 506  
|||||  
DB 458 GGGGAGAGCCCCCACCCTACCGAGGCCCCCTGCACCTCCAGCT 501  
|||||

Search completed: March 18, 2003, 07:55:51  
Job time : 2101 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 07:12:45 ; Search time 210 Seconds

(without alignments)  
9233.193 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861  
Sequence: 1 atgcacgcgtgatgaggggtc.....aacagaagacaccccttc 861

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N.Geneseq-101002.\*

```

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

```

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	861	100.0	1321	24	ABK12137 Human cDNA encodin
2	752	87.3	1140	24	ABK92120 Prostate cancer-as
3	740	85.9	1066	22	AA157868 Human polynucleoti
4	701	81.4	969	21	AAA75151 cDNA encoding a hu
5	701	81.4	1061	21	AAA47429 Sequence encoding
6	696	80.8	969	21	AAA75163 cDNA clone encodin
7	650	75.5	969	21	AAA75164 cDNA clone encodin
8	650	75.5	969	21	AAA75165 cDNA clone encodin
9	401	46.6	408	22	AAF65983 Novel human polynu

C	10	346	40.2	1069	22	AA159654 Human polynucleoti
C	11	302	35.1	693	20	ABK12143 Human MIV-1 homol
	12	271	31.5	812	24	AA252964 Human prostate tum
	13	229	26.6	254	21	AAA41265 Human secreted exp
	14	106	12.3	1879	23	AA584503 DNA encoding novel
	15	51	5.9	51	22	AAH89714 Human coding sequ
	16	49	5.7	60	24	ABN40872 Human spliced tran
	17	44	5.1	878	24	ABK12142 Mouse cDNA encodin
	18	44	5.1	1713	21	AAA75152 cDNA encoding a mu
	19	44	5.1	1713	21	AAA75166 cDNA clone encodin
	20	44	5.1	1713	21	AAA75167 cDNA clone encodin
	21	44	5.1	1713	21	AAA75168 cDNA clone encodin
	22	41	4.8	475	24	ABK12144 Human MIV-1 homol
	23	32	3.7	577	22	ABA50270 Human breast cell
	24	32	3.7	577	22	ABA68210 Human foetal liver
	25	32	3.7	577	22	ABA35222 Probe #13688 for g
	26	32	3.7	577	22	AAK16592 Human bone marrow
	27	32	3.7	577	22	AAK42346 Human bone marrow
	28	32	3.7	577	22	AAI23118 Probe #13051 for g
	29	32	3.7	577	22	AAI48425 Probe #17111 used
	30	32	3.7	577	22	AAI08772 Human genome-deriv
	31	32	3.7	577	24	ABK16402 Human breast cell
	32	32	3.7	1964	22	ABA45134 Human foetal liver
	33	32	3.7	1964	22	ABA55612 Human foetal liver
	34	32	3.7	1964	22	ABA25310 Human bone marrow
	35	32	3.7	1964	22	AAK03841 Human bone marrow
	36	32	3.7	1964	22	AAK29307 Probe #3834 for ge
	37	32	3.7	1964	22	AAI13901 Probe #3953 used t
	38	32	3.7	1964	22	AAI03769 Probe #3760 used t
	39	32	3.7	1964	22	AAI03769 Human genome-deriv
	40	32	3.7	1964	24	ABS03866 Genomic sequence #
C	41	31	3.6	10437	22	AA541709 Human nervous syst
	42	31	3.6	12620	22	ABK16691 DNA encoding novel
C	43	30	3.5	2109	23	AA581477 Human breast cell
	44	29	3.4	401	22	ABA45299 Human foetal liver
	45	29	3.4	401	22	ABA55787 Human foetal liver

## ALIGNMENTS

RESULT 1	ABK12137	standard; cDNA; 1321 bp.
ID	ABK12137	
AC	ABK12137	
XX		
DT	05-JUN-2002	(first entry)
XX		
DE	Human cDNA encoding Mechanically Induced Vascular Receptor 1, MIVR-1.	
XX		
KW	Human; ss; gene; MIVR-1; Mechanically Induced Vascular Receptor 1;	
KW	cytostatic; cardiact; cerebroprotective; antiarteriosclerotic;	
KW	cardiac cell; anti-apoptotic; vascular endothelial cell;	
KW	cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;	
KW	heart failure.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	413..1276
FT		/*tag= a
FT		/product= "MIVR-1"
FT		/note= "This region is specifically claimed in claim 3"
XX		
XX	MO200216416-A2.	
XX		
PN	28-FEB-2002.	
XX		
PD		
XX		
PF	21-AUG-2001; 2001WO-US26089.	
XX		
PR	22-AUG-2000; 2000US-227159P.	
XX		



RESULT 3	
AA157868	
ID	AA157868 standard; cDNA; 1066 BP.
XX	
AC	AA157868;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polynucleotide SEQ ID NO 71.
XX	
XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

NW	Chemokine; drug screening; arthritis; inflammation.
KM	leukaemia; ss.
XX	
OS	
XX	Homo sapiens.
XX	

XX	26-JUL-2001.
PD	
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSEQ-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI: 2001-442253/47.
DR	P-PSDB; AAM38712.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
XX	
PS	Claim 1; SEQ ID NO 71; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	activation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SQ	Sequence 1066 BP; 225 A; 348 C; 325 G; 168 T; 0 other:
XX	
Query Match	85.9%; Score 740; DB 22; Length 1066;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 740; Conservative	0; Mismatches 0; Indels 0; Gaps 0;



Db 130 GCCAGGGGGGAGAGAGAGATGCTGCTCTCAGAGAGATGCTGCTGGCCCTCGGAGA 189  
QY 290 GCACAGTGTAGGACGAAATCCAGAGCCGAGGTCTACGCCCCGCTCGGCCACCG 349  
Db 190 GCACAGTGTAGGACGAAATCCAGAGCCGAGGTCTACGCCCCGCTCGGCCACCG 249  
QY 350 ACCGCGCGGGCGTGGCCCTTGGCCCGAGGGGAGGCTTCGACCCGTTTCAGGCCACT 409  
Db 250 ACCGCGCGGGCGTGGCCCTTGGCCCGAGGGGAGGCTTCGACCCGTTTCAGGCCACT 309  
QY 410 ATCCGTAAGTGCAGCAGAGATGACCTGGCCACCACCATCTGCTCTCAGAGCGGAGG 469  
Db 310 ATCCGTAAGTGCAGCAGAGATGACCTGGCCACCACCATCTGCTCTCAGAGCGGAGG 369  
QY 470 AGCCCCCAGCTTACAGAGGCCCCCTGCACCTCCAGCTTGGGGAGCCCCGAGCAGAGCTGG 529  
Db 370 AGCCCCCAGCTTACAGAGGCCCCCTGCACCTCCAGCTTGGGGAGCCCCGAGCAGAGCTGG 429  
QY 530 AACTGAACGGGAGTGGTGGCGGACACCCCAAAACAGAACCATCTTGCAGAGTACCTGA 589  
Db 430 AACTGAACGGGAGTGGTGGCGGACACCCCAAAACAGAACCATCTTGCAGAGTACCTGA 489  
QY 590 TGGATAGTGCAGCAGCTGGGGCGCCCTGCCCCCAGCAGTAACTCGGCGATCAGCGCCA 649  
Db 490 TGGATAGTGCAGCAGCTGGGGCGCCCTGCCCCCAGCAGTAACTCGGCGATCAGCGCCA 549  
QY 650 CGTGTACAGCAGCGCGGCGGCGCATGGAGGGGCGCCGCCACCTACAGCGAGGTCAATCG 709  
Db 550 CGTGTACAGCAGCGCGGCGGCGCATGGAGGGGCGCCGCCACCTACAGCGAGGTCAATCG 609  
QY 710 GCCACTACCGGGGGTCTCTCTTCAGAGCAGAGAGCAGTGGGGCGCCCTCTCTTGGTGG 769  
Db 610 GCCACTACCGGGGGTCTCTCTTCAGAGCAGAGAGCAGTGGGGCGCCCTCTCTTGGTGG 669  
QY 770 AGGGGAGCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCTGAGACA 829  
Db 670 AGGGGAGCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCTGAGACA 729  
QY 830 AAGGAAGGATTAAGAGAAAGACACCCCTCTC 861  
Db 730 AAGGAAGGATTAAGAGAAAGACACCCCTCTC 761

RESULT 5  
AAAA7429  
ID AAA47429 standard; DNA: 1061 BP.  
AC AAA47429;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Sequence encoding human neuron-associated protein.  
XX  
KW Neuron associated protein; NEUP; neurological disorder; epilepsy;  
KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
KW Alzheimer's disease; Pick's disease; Huntington's disease;  
KW dementia; Parkinson's disease; demyelinating disease; meningitis;  
KW prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;  
KW cerebral palsy; muscular dystrophy; central nervous system; CNS;  
KW peripheral nervous system; PNS; myopathy; schizophrenia;  
KW acanthosis; keratosis; arteriosclerosis; atherosclerosis; bursitis;  
KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;  
KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
KW AIDS; Addison's disease; adult respiratory distress syndrome;  
KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
KW Werner syndrome; trauma; human; ds.  
XX  
OS Homo sapiens.  
XX  
FH key Location/Qualifiers  
FT CDS 101..859  
/\*tag= a

FT /product= Neuron associated protein  
XX  
PN W0200034477-A2.  
XX  
PD 15-JUN-2000.  
XX  
PF 10-DEC-1999; 99NM-US30408.  
XX  
PR 11-DEC-1998; 98US-0210083.  
PR 11-DEC-1998; 98US-9123456.  
PR 09-FEB-1999; 99US-0119365.  
PR 16-MAR-1999; 99US-0124687.  
XX  
PA (INCY-) INCYTE PHARM INC.  
XX  
PI Tang YT, Yue H, Baughn MR, Hillman JL, Lai P, Au-Young J, Yang J;  
PI Lu DAM, Azimzal Y;  
XX  
DR WPI: 2000-423423/36.  
DR P-PSDB; AAB01388.  
XX  
PS Claim 9; Page 136; 145pp; English.  
XX  
CC Human neuron-associated proteins (NEUP) can be used for  
CC treating or preventing a disorder associated with decreased  
CC expression or activity of NEUP. Antagonists of NEUP are useful for  
CC treating or preventing disorder associated with increased expression  
CC or activity of NEUP. NEUP or their fragments or derivatives are  
CC useful for treating neurological disorder such as epilepsy, ischemic  
CC cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's  
CC disease, Pick's disease, Huntington's disease, dementia and  
CC Parkinson's disease. NEUPs are also useful for treating other  
CC demyelinating diseases, bacterial and viral meningitis, prion  
CC diseases including kuru, Creutzfeldt-Jakob disease, nutritional and  
CC metabolic diseases of the nervous system, neurofibromatosis, other  
CC developmental disorders of the central nervous system, cerebral  
CC palsy, neuroskeletal disorders, autonomic nervous system disorders,  
CC cranial nerve disorders, spinal cord diseases, muscular dystrophy and  
CC other neuromuscular disorders, peripheral nervous system disorders,  
CC inherited, metabolic, endocrine, and toxic myopathies, mental  
CC disorders including mood, anxiety and schizophrenic disorders, a cell  
CC proliferative disorder such as actinic keratosis, arteriosclerosis,  
CC atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective  
CC tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal  
CC haemoglobinuria, cancers of the adrenal gland, bladder, bone,  
CC bone marrow, brain, breast, cervix, and an autoimmune/inflammatory  
CC disorder such as acquired immunodeficiency syndrome (AIDS), Addison's  
CC disease, adult respiratory distress syndrome, allergies, ankylosing  
CC spondylitis, amyloidosis, anemia, asthma, Werner syndrome,  
CC complications of cancer, hemodialysis, and extracorporeal circulation,  
CC viral, bacterial, fungal parasitic, protozoal, and helminthic  
CC infections, and trauma. This sequence was given the Incyte ID no.  
CC 1871288CB1.  
XX  
SQ Sequence 1061 BP; 225 A; 342 C; 326 G; 168 T; 0 other;  
XX

Query Match 81.4%; Score 701; DB 21; Length 1061;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 751; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 110 CGGAGCTGAGGTTTGTTCAGATCATCATCATCTGTTGGTATGATGTTGGTGG 169  
Db 105 CGGAGCTGAGGTTTGTTCAGATCATCATCATCTGTTGGTATGATGTTGGTGG 164  
QY 170 TGATCAGTGTGCTGTAGGACACACAGCTGTGCAGGCTCTTATCAGCGGCACA 229  
Db 165 TGATCAGTGTGCTGTAGGACACACAGCTGTGCAGGCTCTTATCAGCGGCACA 224  
QY 230 GCCAGGGGGGAGAGAGAGATGCTGCTCTCAGAGAGATGCTGCTGGCCCTCGGAGA 289

```

Db 225 GCCAGGGGCGAGAGAGAGATGCTGCTCAGAGAGATGCTGCGCCCTGGAGAGA 284
QY 290 GCACAGTGTAGAGCAGAGATCCAGAGCGAGCTCTACGCCGCCCTCGGCCACCG 349
Db 285 GCACAGTGTAGAGCAGAGATCCAGAGCGAGCTCTACGCCGCCCTCGGCCACCG 344
QY 350 ACCGCTGCGCGCTGCGCCCTTCGCCAGCGGAGCGCTTCACACCGCTTCAGCCACT 409
Db 345 ACCGCTGCGCGCTGCGCCCTTCGCCAGCGGAGCGCTTCACACCGCTTCAGCCACT 404
QY 410 ATCCGTAACCTGACAGCAGATGACCTGCAACCCACATCTGCTGTACAGCGGAGG 469
Db 405 ATCCGTAACCTGACAGCAGATGACCTGCAACCCACATCTGCTGTACAGCGGAGG 464
QY 470 AGCCCCACCTTACAGAGGCGCCCTGCAACCTGCGAGCTGGGAGCGCCGACACACTG 529
Db 465 AGCCCCACCTTACAGAGGCGCCCTGCAACCTGCGAGCTGGGAGCGCCGACACACTG 524
QY 530 AACTGAACCGGAGTGGTGGCGGCAACCCCAAGAACATCTTGCAGTGAACCTGA 589
Db 525 AACTGAACCGGAGTGGTGGCGGCAACCCCAAGAACATCTTGCAGTGAACCTGA 584
QY 590 TGGATGTGTGACAGCTGGGCGCGCCCTGCGCCACAGTAACTGGGCAATCAGCGCA 649
Db 585 TGGATGTGTGACAGCTGGGCGCGCCCTGCGCCACAGTAACTGGGCAATCAGCGCA 644
QY 650 CGTGCTACGCGACGCGCGCGCGCATGAGAGGCGCGCGCCACCTACAGGAGGTATG 709
Db 645 CGTGCTACGCGACGCGCGCGCGCATGAGAGGCGCGCGCCACCTACAGGAGGTATG 704
QY 710 GCCACTACCCGGGGTCTCTCTTCACAGCAGCAGAGAGTGGCGCGCTCTCTCTG 769
Db 705 GCCACTACCCGGGGTCTCTCTTCACAGCAGCAGAGAGTGGCGCGCTCTCTCTG 764
QY 770 AGGGGACCGCGCTCCACACACACACATGCGCGCCCTAAGAGCGGAGCATCTGAGAG 829
Db 765 AGGGGACCGCGCTCCACACACACACATGCGCGCCCTAAGAGCGGAGCATCTGAGAG 824
QY 830 AAGAGAGAGATTAACAGAAAGACACCTCTC 861
Db 825 AAGAGAGAGATTAACAGAAAGACACCTCTC 856

RESULT 6
AAA75163
ID AAA75163 standard; cDNA; 969 BP.
XX
XX AAA75163;
XX
XX 15-JAN-2001 (first entry)
XX
XX cDNA clone encoding a human TANGO 261 polypeptide.
DE
XX
XX TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;
XX cellular proliferation; cellular differentiation; cellular adhesion;
XX von Willebrand factor-associated disorder; cell trafficking; cancer;
XX hematopoietic associated disease; atelectasis; pulmonary congestion;
XX oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
XX intestinal disorder; spleen associated disease; renal disorder;
XX cardiovascular disorder; ischemic heart disease; hydrocephalus;
XX brain herniation; iatrogenic disease; inflammation; meningitis;
XX Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;
XX multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 6..764
XX FT /*cag= a
XX FT /product= "TANGO 261"
XX
XX W0200052022-A1.

```

```

XX
PD 08-SEP-2000.
XX
PF 01-MAR-2000; 2000WO-US05226.
XX
PR 01-MAR-1999; 99US-0122458.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CJ;
XX
PI WPI: 2000-579269/54.
XX
PT P-PSDB: AAB18461.
XX
PT Novel human and murine secreted proteins designated TANGO 216, 261,
XX 262, 266 and 267 useful as modulating agents of cellular processes,
XX e.g. for treating cancer.
XX
PS Disclosure; Page -: 175pp; English.
XX
CC AAA75163-65 encode human TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis,
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC bacterial and viral meningitis, Alzheimer's disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.
XX
XX Sequence 969 BP: 211 A; 317 C; 293 G; 148 T; 0 other:
XX
XX
XX Query Match 80.8%; Score 696; DB 21; Length 969;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 746; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 115 CTGAGTTTGTTCAGATCATCATCATGCTGGTGTGATGATGATGATGATGATGATC 174
Db 15 CTGAGTTTGTTCAGATCATCATCATGCTGGTGTGATGATGATGATGATGATGATC 74
QY 175 AGTGCTCTGTGACCACTACAACTCTGTGACAGGCTCTTCATACACCGGACAGCAG 234
Db 75 AGTGCTCTGTGACCACTACAACTCTGTGACAGGCTCTTCATACACCGGACAGCAG 134
QY 235 GGGGAGGAGAGAAATGCGCTGTCTCTCAAGAGATGCGCTGGCGCTGGAGAGCACA 294
Db 135 GGGGAGGAGAGAAATGCGCTGTCTCTCAAGAGATGCGCTGGCGCTGGAGAGCACA 194
QY 295 GTGTGAGCAAGGAATCCAGAGCGCAGGTACGCGCCGCTGGCGCCACGACCGC 354
Db 195 GTGTGAGCAAGGAATCCAGAGCGCAGGTACGCGCCGCTGGCGCCACGACCGC 254
QY 355 CTGGCGCTGGCGCCCTTCGCCAGCGGAGCGCTTCACCGCTTCAGGCCACTATCGG 414
Db 255 CTGGCGCTGGCGCCCTTCGCCAGCGGAGCGCTTCACCGCTTCAGGCCACTATCGG 314
QY 415 TACCTGACGAGAGATGACCTGCGCACCCACATCGCTGTCAACGGGAGGAGCGCC 474

```

RESULT 7  
AAA75164  
ID AAA75164 standard; cDNA; 969 BP.

KW TANGO 26; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; aelcelsais; pulmonary congestion;  
 KW Oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
 XX Homo sapiens.

PN W0200052022-A1.

PD 08-SEP-2000.

PF 01-MAR-2000; 2000WO-US05226.

PR 01-MAR-1999; 99US-0122458.

PA (MILL-) MILLENNIUM PHARM INC.

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

DR WPI; 2000-579269/54.

DR P-PSDB; AAB18462.

CC AAA7516-65 encode human TANGO 261 proteins. The specification also  
CC describes TANGO 265, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
CC polypeptides can be used to modulate cellular proliferation, modulate  
CC cellular differentiation and/or modulate cellular adhesion. The  
CC proteins can be used to treat any von Willebrand factor-associated  
CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
CC and cell trafficking and/or migration, modulate cellular interactions,  
CC modulate cell adhesion in proliferative disorders, such as cancer,  
CC modulate the proliferation, differentiation, and/or function of cells  
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
CC and hematopoietic associated diseases and disorders, atelectasis,  
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
CC asthma and bronchiectasis, intestinal disorders, spleen associated  
CC diseases, modulate renal disorders, treat cardiovascular disorders such  
CC as ischemic heart disease, modulate the proliferation, differentiation,  
CC and/or function of bone and cartilage cells and to treat bone and/or  
CC cartilage associated diseases or disorder. They may also be used to  
CC treat disorders associated with the ovaries, and cerebral oedema,  
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
CC bacterial and viral meningitis, Alzheimer's Disease, cerebral  
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
CC hydrocephalus and encephalitis, and treat hepatic disorders.  
CC note: The present sequence does not appear in the specification; it was  
CC created using information provided.

SQ Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other;

Query Match	75.5%	Score 650;	DB 21;	Length 969;
Best Local Similarity	99.7%	Pred. No. 5.8e-295;		
Matches 750; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

[illegible]









|||||  
Db 467 ATAGTCCAGAGGCTGGCGGCCCTGCCCCAGAGTAATCTGGGATCAGCGCCACGT 408  
OY 653 GCTACGCGACCGGGGGCGATGAGGGGCGCCGCCACCTACAGAGATATGGGCC 712  
Db 407 GCTACGCGACCGGGGGCGATGAGGGGCGCCGCCACCTACAGAGATATGGGCC 348  
OY 713 ACTACCGGGGGCTCTCTTCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 772  
Db 347 ACTACCGGGGGCTCTCTTCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288  
OY 773 GG 774  
Db 287 GG 286

RESULT 12

AA252964  
ID AA252964 standard; cDNA; 812 BP.  
XX  
AC AA252964;

14-MAR-2000 (first entry)

Human prostate tumor cDNA library derived EST fragment #107.  
XX  
XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;  
XX treatment; ds.

Homo sapiens.

DE19820190-A1.

04-NOV-1999.

28-APR-1998; 98DE-1020190.

28-APR-1998; 98DE-1020190.

(META-) METAGEN GES GENOMFORSCHUNG MBH.

Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

WPI: 1999-621386/54.

P-PSDB; AAY74135, AAY74136, AAY74137.

PT New human nucleic acid sequences from pancreatic tumors, and related  
XX proteins

Claim 2; Page 269-270; 502pp; German.

CC This invention describes novel polypeptides and their encoding nucleic  
CC acids derived from human pancreatic tumor tissue which have cytostatic  
CC activity. The sequences are also useful in producing pharmaceutical  
CC compositions for treatment of pancreatic tumors. AA252858-253014  
CC represent expressed sequence tag (EST) fragments derived from a human  
CC pancreatic tumor cDNA library and which encode the proteins represented  
CC in AAY73814-Y74252.

SQ Sequence 812 BP; 157 A; 272 C; 237 G; 146 T; 0 other;

Query Match 31.5%; Score 271; DB 20; Length 812;  
Best Local Similarity 99.7%; Pred. No. 3.7e-117;  
Matches 331; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 173 TCACGTGCTGCTGACCACTACAGCTGTCTCCACGCTCTTCATCAGCGGACAGCC 232  
Db 426 TCACGTGCTGCTGACCACTACAGCTGTCTCCACGCTCTTCATCAGCGGACAGCC 485  
OY 233 AGGGGCGGAGAGAGAGATGCTCTCTCAAGAGATGCTGTGGCCCTCGAGAGCA 292  
Db 486 AGGGGCGGAGAGAGAGATGCTCTCTCAAGAGATGCTGTGGCCCTCGAGAGCA 545

OY 293 CAGTGTACGCAACGAGATCCAGAGCGGAGTCTACGCGCCCGCTCGGCCACGAGC 352  
Db 546 CAGTGTACGCAACGAGATCCAGAGCGGAGTCTACGCGCCCGCTCGGCCACGAGC 605  
OY 353 GCCTGGCGGTGCGGCCCTTCGCGGAGGAGCGGCTTCACACCGCTTCACACCTATC 412  
Db 606 GCCTGGCGGTGCGGCCCTTCGCGGAGGAGCGGCTTCACACCGCTTCACACCTATC 665  
OY 413 CGTACCTGACGACAGATATGACCTGACACCCACATCTGCTGTACAGAGGGAGAGC 472  
Db 666 CGTACCTGACGACAGATATGACCTGACACCCACATCTGCTGTACAGAGGGAGAGC 725  
OY 473 CCCACCTACACAGGCGCCCTG 494  
Db 726 CCCACCTACACAGGCGCCCTG 747

RESULT 13

AAA41265  
ID AAA41265 standard; cDNA; 254 BP.  
XX  
AC AAA41265;

21-AUG-2000 (first entry)

Human secreted expressed sequence tag SEQ ID NO:5.

XX Human; mouse; xenopus; rat; secreted expressed sequence tag; SEST;  
XX expressed sequence tag; EST; probe; chemokine; proliferative;  
XX immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;  
XX thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;  
XX antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;  
XX antitumor; osteoprotective; neuroprotective; nocitropic; antiproliferative;  
XX cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
XX vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
XX insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
XX central nervous system disorder; Alzheimer's disease; stroke;  
XX Parkinson's disease; Huntington's disease; coagulation disorder;  
XX haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
XX tumour; infection; depression; psoriasis; ss.

Homo sapiens.

WO200021990-A1.

20-APR-2000.

15-OCT-1999; 99MO-US24205.

15-OCT-1998; 98US-0104435.

(GEM) GENETICS INST INC.

Jacobs R, McCoy JM, Lavaille ER, Collins-Racie LA, Evans C;

Merberg D, Tracy M;

WPI: 2000-317937/27.

PT Isolated polynucleotides, and encoded proteins, comprising secreted  
XX expressed sequence tags (SESTs), useful for treating various disorders  
XX such as autoimmune, infectious, and central nervous system disorders -  
XX Claim 1; Page 180; 618pp; English.

CC AAA41261 to AAA43419 represent specifically claimed secreted expressed  
CC sequence tags (SESTs). Isolated from human, mouse, xenopus and rat  
CC tissue sources. The SESTs can have a range of activities depending on  
CC the tissues they were isolated from. The activities include:  
CC chemokine; proliferative; immunomodulatory; haematopoietic;  
CC chemokine; analgesic; haemostatic; thrombolytic; antiinflammatory;  
CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;  
CC antiasthmatic; vulnary; antitumor; osteoprotective; neuroprotective;

CC neotropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The SESTs can be used for gene  
 CC therapy and in vaccines. The SESTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the SESTs. Proteins encoded by the SESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given  
 CC in the exemplification of the present invention.

CC Sequence 254 BP; 46 A; 85 C; 82 G; 41 T; 0 other;

Query Match 26.6%; Score 229; DB 21; Length 254;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-97;

Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 151 ATGATGGTGTGATGCTGTGATCAGTCTGCTGAGCCACTACAGCTGTCTGACGG 210  
 CC |  
 CC 21 ATGATGGTGTGATGCTGTGATCAGTCTGCTGAGCCACTACAGCTGTCTGACGG 80  
 CC |  
 CC 211 TCCTTCATCAGCCGCGACAGCCAGGGGGGAGAGAAAGATGCCCTGTCTCAGAGA 270  
 CC |  
 CC 81 TCCTTCATCAGCCGCGACAGCCAGGGGGGAGAGAAAGATGCCCTGTCTCAGAGA 140  
 CC |  
 CC 271 TCCTGTGTGCTTCGAGAGACAGTGTCTCAGCAACGATCCAGAGCCGAGGCTCTAC 330  
 CC |  
 CC 141 TGCTGTGTGCTTCGAGAGACAGTGTCTCAGCAACGATCCAGAGCCGAGGCTCTAC 200  
 CC |  
 CC 331 GCCCGGCTCGGCGCCACGACCGCTGCGCCGCGCCCTTGCCCAAGC 379  
 CC |  
 CC 201 GCCCGGCTCGGCGCCACGACCGCTGCGCCGCGCCCTTGCCCAAGC 249  
 CC |

#### RESULT 14

ID AAS84503 standard; cDNA: 1879 BP.

AC AAS84503;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #20307.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX MO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dzmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX P-PSDB; ABG20316.

PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess

PT biodiversity -  
 XX Claim 1; SEQ ID No 20307; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1879 BP; 488 A; 491 C; 502 G; 398 T; 0 other;

Query Match 12.3%; Score 106; DB 23; Length 1879;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-40;

Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 376 CAGCGGGAGCGCTTCCACCGCTTCCAGCCACTTCCGTACTGACAGCAGATCGAC 435  
 CC |  
 CC 333 CAGCGGGAGCGCTTCCACCGCTTCCAGCCACTTCCGTACTGACAGCAGATCGAC 392  
 CC |  
 CC 436 CTGCCACCCACCATCTCGCTGTCAAGCGGAGAGAGCCGCCACCGT 481  
 CC |  
 CC 393 CTGCCACCCACCATCTCGCTGTCAAGCGGAGAGAGCCGCCACCGT 438  
 CC |

#### RESULT 15

ID AAH89714 standard; DNA; 51 BP.

AC AAH89714;

DT 01-OCT-2001 (first entry)

DE Human coding sequence polymorphic site SEQ ID NO: 495.

XX Human; single nucleotide polymorphism; SNP; paternity test;  
 KW forensic test; aberrant protein expression; ds.

OS Homo sapiens.

XX MO200151670-A2.

PD 19-JUL-2001.

XX 05-JAN-2001; 2001WO-US00322.

XX 07-JAN-2000; 2000US-0174962.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach MD;

XX WPI: 2001-451871/48.

XX P-PSDB; AAM00597.

PT Isolated human polynucleotides containing single nucleotide  
 PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,

PT infection and diabetes -

PS Claim 1; Page 246; 475pp; English.

CC The present invention relates to human nucleic acids containing single  
CC nucleotide polymorphisms (SNPs). These can be used in forensic and  
CC paternity tests, and to aid in the treatment of diseases associated with  
CC aberrant protein expression, including cancer, amyloidosis, diabetes,  
CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,  
CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,  
CC meningitis, muscular disorders, dementia, neurological diseases, tuberculous  
CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,  
CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or  
CC autoimmunity. The present sequence is a polymorphism-containing  
CC oligonucleotide fragment of the invention.

SQ Sequence 51 BP; 12 A; 16 C; 18 G; 5 T; 0 other;

Query Match	5.9%;	Score 51;	DB 22;	Length 51;
Post Local Similarity	100.0%;	Post Local	6.6414	

Matches	51;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

206 CACGGTCTTCATCAGCCGGCACAGCCAGGGGGGAGAGAGAATGCC 256

Db 1 CACGGTCTTCATCAGCCGGCACAGCCAGGGGGCGAGAGAGAGATGCCC 51

Search completed: March 18, 2003, 07:20:34  
Job time : 215 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 07:16:10 ; Search time 47 Seconds  
(without alignments)  
5618.057 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861

Sequence: 1 atgcacgccttgatg999gt.....aacagaagacaccccttc 861

Scoring table:

OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata/1/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/lna/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/lna/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/lna/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/lna/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	2.7	377	2	US-08-332-766A-1
2	21	2.4	633	1	US-08-234-783-1
3	21	2.4	633	1	US-08-456-907-1
4	21	2.4	633	5	PCT-US95-05523-1
5	21	2.4	1327	1	US-08-165-315D-3
6	21	2.4	4659	4	US-08-264-578-10
7	21	2.4	8147	4	US-09-514-247A-9
8	21	2.4	8252	1	US-08-046-585-15
9	21	2.4	8252	1	US-08-393-703-15
10	21	2.4	8252	5	PCT-US93-11721-15
11	20	2.3	50	4	US-08-753-247-22
12	20	2.3	51	4	US-08-753-247-23
13	20	2.3	450	4	US-09-370-838-145
14	20	2.3	640	2	US-08-835-099A-16
15	20	2.3	640	3	US-09-157-349-16
16	20	2.3	804	2	US-08-835-099A-10
17	20	2.3	804	3	US-09-157-349-10
18	20	2.3	1028	4	US-08-118-200-1
19	20	2.3	1028	4	US-08-458-745-1
20	20	2.3	1044	4	US-07-975-526-3
21	20	2.3	1044	4	US-07-974-409C-425
22	20	2.3	1189	1	US-07-781-034-4
23	20	2.3	1189	5	PCT-US92-08328-4
24	20	2.3	1605	4	US-09-124-541-3
25	20	2.3	2130	4	US-07-952-800-1
26	20	2.3	6733	4	US-09-124-541-2
27	20	2.3	10348	2	US-08-457-273B-41

28	20	2.3	10348	3	US-08-556-419-13	Sequence 13, Appl
29	20	2.3	10348	4	US-09-041-886-14	Sequence 14, Appl
30	20	2.3	10366	1	US-08-246-982A-5	Sequence 5, Appl
31	20	2.3	10366	1	US-08-453-265-5	Sequence 27, Appl
32	19	2.2	49	1	US-08-155-171B-27	Sequence 28, Appl
33	19	2.2	49	1	US-08-155-171B-28	Sequence 27, Appl
34	19	2.2	49	2	US-08-435-998-27	Sequence 28, Appl
35	19	2.2	49	2	US-08-435-998-28	Sequence 28, Appl
36	19	2.2	49	4	US-09-813-781-69	Sequence 36, Appl
37	19	2.2	53	1	US-08-155-171B-36	Sequence 36, Appl
38	19	2.2	53	2	US-08-435-998-36	Sequence 36, Appl
39	19	2.2	81	4	US-09-497-933A-20	Sequence 129, App
40	19	2.2	91	1	US-08-142-551B-129	Sequence 130, App
41	19	2.2	91	1	US-08-142-551B-130	Sequence 30, Appl
42	19	2.2	276	2	US-08-332-766A-30	Sequence 328, App
43	19	2.2	440	4	US-09-397-787-328	Sequence 307, App
44	19	2.2	491	4	US-09-643-597-307	Sequence 116, App
45	19	2.2	649	4	US-08-998-416-116	

#### ALIGNMENTS

```
RESULT 1
US-08-332-766A-1
; Sequence 1, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFERYS, Alec J.
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-332-766A-1
;
Query Match      2.7%; Score 23; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 143 TGGTGTGATGATGATGATGATGATG 165

Db 123 TGGTGTGATGATGTGATGCTG 145

## RESULT 2

US-08-234-783-1  
Sequence 1, Application US/08234783  
Patent No. 5622835  
GENERAL INFORMATION:  
APPLICANT: Herlyn, Meenhard  
APPLICANT: Morris, Jennifer  
APPLICANT: Rauscher III, Frank J.  
APPLICANT: Rodeck, Ulrich  
TITLE OF INVENTION: Wt1 Monoclonal Antibodies and Methods of  
TITLE OF INVENTION: Use Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/234,783  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST48USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 633 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..630

US-08-234-783-1

Query Match 2.4%; Score 21; DB 1; Length 633;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GCACCGCGCGCGCGCGCGCG 46  
Db 208 GCACCGCGCGCGCGCGCGCG 228

## RESULT 3

US-08-456-907-1  
Sequence 1, Application US/08456907  
Patent No. 5633142  
GENERAL INFORMATION:  
APPLICANT: Herlyn, Meenhard  
APPLICANT: Morris, Jennifer  
APPLICANT: Rauscher III, Frank J.  
APPLICANT: Rodeck, Ulrich  
TITLE OF INVENTION: Wt1 Monoclonal Antibodies and Methods of  
TITLE OF INVENTION: Use Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,907  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: US 08/234,783  
APPLICATION NUMBER: 28-APR-1994  
FILING DATE: 28-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST48AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 633 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..630

US-08-456-907-1

Query Match 2.4%; Score 21; DB 1; Length 633;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GCACCGCGCGCGCGCGCGCG 46  
Db 208 GCACCGCGCGCGCGCGCGCG 228

## RESULT 4

PCT-US95-05523-1  
Sequence 1, Application PC/TUS9505523  
GENERAL INFORMATION:  
APPLICANT: The Wistar Institute of Anatomy and Biology  
TITLE OF INVENTION: Wt1 Monoclonal Antibodies and  
TITLE OF INVENTION: Methods of Use Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05523  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/234,783  
FILING DATE: 28-APR-1994



```
ATTORNEY/AGENT INFORMATION:
: NAME: Bak, Mary E.
: REGISTRATION NUMBER: 31,215
: REFERENCE/DOCKET NUMBER: WST48PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-540-9200
: TELEFAX: 215-540-5818
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 633 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..630
: PCT-US95-05523-1

Query Match          2.4%: Score 21; DB 5; Length 633;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GCACCGCGCGCGCGCGCG 46
Db 208 GCACCGCGCGCGCGCGCG 228

RESULT 5
US-08-165-315D-3
: Sequence 3, Application US/08165315D
: Patent No. 5525716
: GENERAL INFORMATION:
: APPLICANT: Odd-Arne Olsen
: APPLICANT: Roger Kalla
: TITLE OF INVENTION: Promoter
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: St. Onge, Steward, Johnston & Reens
: STREET: 986 Bedford Street
: CITY: Stamford
: STATE: Connecticut
: COUNTRY: U.S.A.
: ZIP: 06905
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" 1.44 Mb diskette
: COMPUTER: IBM PC
: OPERATING SYSTEM: MS DOS
: SOFTWARE: Word Processor
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/165,315D
: FILING DATE: 10 December 1993
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9324707.0
: FILING DATE: 2 December 1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Mary M. Krinsky
: REGISTRATION NUMBER: 32423
: REFERENCE/DOCKET NUMBER: 2105-P0001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-324-6155
: TELEFAX: 201-327-1096
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1327
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: DESCRIPTION: gene
: FRAGMENT TYPE: gene
: FEATURE:
```

```
NAME/KEY: Ltp2 gene
US-08-165-315D-3

Query Match          2.4%: Score 21; DB 1; Length 1327;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 GTGATGATGATGATGATG 168
Db 918 GTGATGATGATGATGATG 938

RESULT 6
US-08-264-578-10
: Sequence 10, Application US/08264578
: Patent No. 6391566
: GENERAL INFORMATION:
: APPLICANT: FOLDES, Robert L.
: APPLICANT: ADAMS, Sally-Lin
: APPLICANT: KAMROU, Rajender
: APPLICANT: DUNCAN, H. Scott
: TITLE OF INVENTION: Modulatory proteins of Human CNS
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington, D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/264,578
: FILING DATE: 23-JUN-1994
: CLASSIFICATION: 436
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/987,953
: FILING DATE: 11-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 16777/261/ALLE
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4659 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1099..3753
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: 1099..1152
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 1153..3753
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 2781..2838
: OTHER INFORMATION: /function= "transmembrane domain"
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 2895..2958
: OTHER INFORMATION: /function= "transmembrane domain"
```

RESULT 9  
 US-08-393-703-15/c  
 Sequence 15. Application US/08393703  
 Patent No. 5585239  
 GENERAL INFORMATION:  
 APPLICANT: Lamarco, Kelly  
 APPLICANT: Wilson, Angus  
 APPLICANT: Herr, Winship  
 TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN  
 TITLE OF INVENTION: HOST CELL FACTOR  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: FLIERL, HOBACH, TEST, ALBRITTON & HERBERT  
 STREET: 4 Embarcadero Center, Suite 3400  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94111-4187  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/393,703  
 FILING DATE: 24-FEB-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Osman, Richard A  
 REGISTRATION NUMBER: 36,627  
 REFERENCE/DOCKET NUMBER: A-57503-2/RAO  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 781-1989  
 TELEFAX: (415) 398-3249  
 TELEX: 910 277299  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8252 base pairs

TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-393-703-15

Query Match 2.4%; Score 21; DB 1; Length 8252;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 TGTGTGATGATGTGTATGG 163  
|||||  
Db 2748 TGTGTGATGATGTGTATGG 2728

RESULT 10  
PCT-US93-11721-15/C  
Sequence 15, Application PC/TUS9311721  
GENERAL INFORMATION:  
APPLICANT: Lamarco, Kelly  
APPLICANT: Wilson, Angus  
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:  
TITLE OF INVENTION: HOST CELL FACTOR  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/11721  
FILING DATE: 03-DEC-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: RP-57503-1/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8252 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
PCT-US93-11721-15

Query Match 2.4%; Score 21; DB 5; Length 8252;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 TGTGTGATGATGTGTATGG 163  
|||||  
Db 2748 TGTGTGATGATGTGTATGG 2728

RESULT 11  
US-08-753-247-22/C  
Sequence 22, Application US/08753247  
Patent No. 6210929  
GENERAL INFORMATION:  
APPLICANT: SCHLOKAT, Uwe

APPLICANT: FISCHER, Bernhard  
APPLICANT: FALKNER, Falko-Guenther  
APPLICANT: DORNER, Friedrich  
APPLICANT: EIBL, Johann  
TITLE OF INVENTION: A FUSION PROTEIN COMPRISING A FURIN  
DERIVATIVE OR A DERIVATIVE OF A FURIN ANALOGUE AND A  
TITLE OF INVENTION: HETEROLOGOUS SEQUENCE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/753,247  
FILING DATE: 22-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AT 1928/95  
FILING DATE: 24-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40433/149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-753-247-22

Query Match 2.3%; Score 20; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 TGATGATGATGTGTGTG 168  
|||||  
Db 42 TGATGATGATGTGTGTG 23

RESULT 12  
US-08-753-247-23  
Sequence 23, Application US/08753247  
Patent No. 6210929  
GENERAL INFORMATION:  
APPLICANT: SCHLOKAT, Uwe  
APPLICANT: FISCHER, Bernhard  
APPLICANT: FALKNER, Falko-Guenther  
APPLICANT: DORNER, Friedrich  
APPLICANT: EIBL, Johann  
TITLE OF INVENTION: A FUSION PROTEIN COMPRISING A FURIN  
DERIVATIVE OR A DERIVATIVE OF A FURIN ANALOGUE AND A  
TITLE OF INVENTION: HETEROLOGOUS SEQUENCE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA

ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/753,247  
FILING DATE: 22-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AT 1928/95  
FILING DATE: 24-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40433/149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 51 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-753-247-23

Query Match 2.3%; Score 20; DB 4; Length 51;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 149 TGATGATGTGATGTGTG 168  
|||||  
DB 13 TGATGATGTGATGTGTG 32

RESULT 13  
US-09-370-838-145/C  
Sequence 145, Application US/09370838  
Patent No. 6444425  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Mohamath, Radooh  
APPLICANT: Secrist, Heather  
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
FILE REFERENCE: 210121.475C1  
CURRENT APPLICATION NUMBER: US/09/370,838  
CURRENT FILING DATE: 1999-08-09  
EARLIER APPLICATION NUMBER: US 09/285,323  
NUMBER OF SEQ ID NOS: 289  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 145  
LENGTH: 450  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-370-838-145

Query Match 2.3%; Score 20; DB 4; Length 450;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 CACCGCGCGCGCGCGCG 46  
|||||  
DB 77 CACCGCGCGCGCGCGCG 58

RESULT 14

US-08-835-099A-16/C  
Sequence 16, Application US/08835099A  
Patent No. 5874277  
GENERAL INFORMATION:  
APPLICANT: SHINTANI, Yasushi  
APPLICANT: NISHI, Kazuori  
APPLICANT: KAWAMOTO, Tomohiro  
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/835,099A  
FILING DATE: 04-APR-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 083649/1996  
FILING DATE: 05-APR-1996  
APPLICATION NUMBER: 97105508.2  
FILING DATE: 03-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 47342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 640 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-835-099A-16

Query Match 2.3%; Score 20; DB 2; Length 640;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 CACCGCGCGCGCGCGCG 46  
|||||  
DB 145 CACCGCGCGCGCGCGCG 126

RESULT 15  
US-09-157-349-16/C  
Sequence 16, Application US/09157349  
Patent No. 6068990  
GENERAL INFORMATION:  
APPLICANT: SHINTANI, Yasushi  
APPLICANT: NISHI, Kazuori  
APPLICANT: KAWAMOTO, Tomohiro  
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA

```

; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 640 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-157-349-16

Query Match 2.3%; Score 20; DB 3; Length 640;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 CACCGCGCGCGCGCGCGCG 46
    |||
Db 145 CACCGCGCGCGCGCGCGCG 126
```

Search completed: March 18, 2003, 08:16:26  
Job time : 80 secs

**THIS PAGE BLANK (USPTO)**



Db	121	TTTGTCAGATCATCATCATCGTGTGGTATATATGGTATGGTGGTGGTGGTATACCTGC	180
Qy	181	CTGCTGAGCCACATCAACAGCTGTTCAGAGGTCTTTCATCAAGCCGGCACAGCCAGGGCGG	240
Db	181	CTGCTGAGCCACATCAACAGCTGTTCAGAGGTCTTTCATCAAGCCGGCACAGCCAGGGCGG	240
Qy	241	AGGAGAGAAGATGGCCCTTCTCTCAAGAAAGATGCTGGCCCTTCGGAGACACAGTCTCA	300
Db	241	AGGAGAGAAGATGGCCCTTCTCTCAAGAAAGATGCTGGCCCTTCGGAGACACAGTCTCA	300
Qy	301	GGCAACGGAAATCCAGAGAGCCGCAAGTCTACGCCGCCGCTCGGCCACAGCAGCCCTGGCC	360
Db	301	GGCAACGGAAATCCAGAGAGCCGCAAGTCTACGCCGCCGCTCGGCCACAGCAGCCCTGGCC	360
Qy	361	GTGGCGCCCTTTCGGCCAGGGGAGCGGCTTCACCGCTTCCAGGCCACCTATCCGTACTG	420
Db	361	GTGGCGCCCTTTCGGCCAGGGGAGCGGCTTCACCGCTTCCAGGCCACCTATCCGTACTG	420
Qy	421	CAGCAGAGATCGACTGCCACCCACCACATCTCGCTGTCAAGCGGGAGAGCCGCCACCC	480
Db	421	CAGCAGAGATCGACTGCCACCCACCACATCTCGCTGTCAAGCGGGAGAGCCGCCACCC	480
Qy	481	TACCAAGGGGCCCCCTGACCCCTTCACCTTCGGGAGCCCCCAGCAGAGCTGGAACTGAACCGG	540
Db	481	TACCAAGGGGCCCCCTGACCCCTTCACCTTCGGGAGCCCCCAGCAGAGCTGGAACTGAACCGG	540
Qy	541	GAGTGGGGTGGCGCGGACCCCAAGAAACATCTTCAGCAGTGCAGCTGATGATGATGGCC	600
Db	541	GAGTGGGGTGGCGCGGACCCCAAGAAACATCTTCAGCAGTGCAGCTGATGATGATGGCC	600
Qy	601	AGGCTGGGGCGGCCCCCTGACCCCAAGCAGTACCTGGGGCATCAAGCGGCACGTGCTACGGC	660
Db	601	AGGCTGGGGCGGCCCCCTGACCCCAAGCAGTACCTGGGGCATCAAGCGGCACGTGCTACGGC	660
Qy	661	AGCGGGGGGCGCATGAGAGGGGCGCGCCCAAGCAGTACAGCAGGATGCATGGGCGACTACCGG	720
Db	661	AGCGGGGGGCGCATGAGAGGGGCGCGCCCAAGCAGTACAGCAGGATGCATGGGCGACTACCGG	720
Qy	721	GGGTCTCTCTTCAGACACCAAGAGGAGTGGGCGGCCCTCTTGTGTGAGAGGGAGCCCGG	780
Db	721	GGGTCTCTCTTCAGACACCAAGAGGAGTGGGCGGCCCTCTTGTGTGAGAGGGAGCCCGG	780
Qy	781	CTTCACCAACACATCGGGGCCCCCTGAGAGGGGCGGCATCTTGAGAGCAAGAGAGAT	840
Db	781	CTTCACCAACACATCGGGGCCCCCTGAGAGGGGCGGCATCTTGAGAGCAAGAGAGAT	840
Qy	841	AAACAGAAAGAGACACCCCTCTC	861
Db	841	AAACAGAAAGAGACACCCCTCTC	861

```

; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (413) ... (1273)
;
US-09-934-249-1

```

Query Match	100.0%;	Score 861;	DB 10;	Length 1321;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 861;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGCAACGGCTGTAATGGGGGGTCAACAGACACCGCCGCCGCCGGCGGGAGGCCCAATGTC	60
Db	413	ATGCAACGGCTGTAATGGGGGGTCAACAGACACCGCCGCCGCCGGCGGGAGGCCCAATGTC	472
QY	61	TCCTGCACGTGCACCTGCAACAGCCTCTTTGTTCAGAGCATGGAGATCAAGAGCTGGAG	120
Db	473	TCCTGCACGTGCACCTGCAACAGCCTCTTTGTTCAGAGCATGGAGATCAAGAGCTGGAG	532
QY	121	TTTGTTCAGATCAATCAATCATCTGTGTGTGTATATATGTGTATGTGTGTATCACTGTC	180
Db	533	TTTGTTCAGATCAATCAATCATCTGTGTGTGTATATATGTGTATGTGTGTATCACTGTC	592
QY	181	CTGTGTGACCACTACAAAGCTGTCTGCACAGGTCCTTGATCAGCCGAGCAGCCAGAGGAGCG	240
Db	593	CTGTGTGACCACTACAAAGCTGTCTGCACAGGTCCTTGATCAGCCGAGCAGCCAGAGGAGCG	652
QY	241	AGGAGAGAAATGATGCTGCTGTCTCAGAAAGATGCTCTTGCCCTCGGAGAGCAAGTGTCA	300
Db	653	AGGAGAGAAATGATGCTGCTGTCTCAGAAAGATGCTCTTGCCCTCGGAGAGCAAGTGTCA	712
QY	301	GGCAACGGAAATCCAGAGCCGCGCAGAGCTAAGCCGCCCTCGGCGCCACGACCGCCTGGGC	360
Db	713	GGCAACGGAAATCCAGAGCCGCGCAGAGCTAAGCCGCCCTCGGCGCCACGACCGCCTGGGC	772
QY	361	GTGCGCCCTTTCGCCACAGCGGGAGAGCGCTTCCACGCGCTTCAGCCCACTATCCGTAACGTG	420
Db	773	GTGCGCGCCCTTTCGCCACAGCGGGAGAGCGCTTCCACGCGCTTCAGCCCACTATCCGTAACGTG	832
QY	421	CAGACAGATATCGACGCGCCACCCACCATCTGTGCTGCAGAGGGGAGAGGCCGCCACCC	480
Db	833	CAGACAGATATCGACGCGCCACCCACCATCTGTGCTGCAGAGGGGAGAGGCCGCCACCC	892
QY	481	TACCAGGGCCCTGCACCCCTCCAGCTTCGSGGAGCCCGAGCAGACAGCTGGAACTGAACCGG	540
Db	893	TACCAGGGCCCTGCACCCCTCCAGCTTCGSGGAGCCCGAGCAGACAGCTGGAACTGAACCGG	952
QY	541	GAGTCGCTGGCGGCACCCCCAAACAAACAAACATCTTGACAGTGAAGCTGATGGATAGTGCC	600
Db	953	GAGTCGCTGGCGGCACCCCCAAACAAACAAACATCTTGACAGTGAAGCTGATGGATAGTGCC	1012
QY	601	AGGCTGGGGCGGCCCTTCGCCCCCAGCAGTAATCTGGGGCATCAGGCGCAAGCTCAAGGC	660
Db	1013	AGGCTGGGGCGGCCCTTCGCCCCCAGCAGTAATCTGGGGCATCAGGCGCAAGCTCAAGGC	1072
QY	661	AGCGGGGGCGCATGAGAGGGGCCGCCCACTACAGCGAGGTATCGGCCACTTAACCG	720
Db	1073	AGCGGGGGCGCATGAGAGGGGCCGCCCACTACAGCGAGGTATCGGCCACTTAACCG	1132
QY	721	GGGTCCTCTTCAGAGCAAGCAGACAGCACTATGGGGCGCCCTCTTCTGTGAAGGAGACCGG	780
Db	1133	GGGTCCTCTTCAGAGCAAGCAGACAGCACTATGGGGCGCCCTCTTCTGTGAAGGAGACCGG	1192
QY	781	CTCCACACACACATCGCGCCCTTAGAGAGCAGCAGCCATCTGGAGCAAAAGAGAGAT	840
Db	1193	CTCCACACACACATCGCGCCCTTAGAGAGCAGCAGCCATCTGGAGCAAAAGAGAGAT	1252
QY	841	AAACAGAAAGAGACCTCTC 861	
Db	1253	AAACAGAAAGAGACCTCTC 1273	

RESULT 3  
US-10-098-841-71







Db 1367 TCCTC 1370

## RESULT 6

```
US-09-934-249-14/c
; Sequence 14, Application US/09934249
; Patent No. US20020115081A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Richard T.
; APPLICANT: Landschulz, Katherine T.
; APPLICANT: Turi, Thomas G.
; APPLICANT: Thompson, John F.
; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS
; FILE REFERENCE: P0738/7001/ERP/KA
; CURRENT APPLICATION NUMBER: US/09/934, 249
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (639)...(639)
; OTHER INFORMATION: a, c, g, or t/u
US-09-934-249-14
```

Query Match 35.1%; Score 302; DB 10; Length 693;

Best Local Similarity 100.0%; Pred. No. 1.4e-135; Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 473 CCCACCTTACAGAGGCGCCCTGACACCCCTTGGGAGCCCGAGCAGCAGCTGAGAC 532
Db 587 CCCACCTTACAGAGGCGCCCTGACACCCCTTGGGAGCCCGAGCAGCAGCTGAGAC 528
QY 533 TGAACCGGAGTCTGTCGCGCCACCCCAACAGAACCATTTGACAGTAGTACCTGATGG 592
Db 527 TGAACCGGAGTCTGTCGCGCCACCCCAACAGAACCATTTGACAGTAGTACCTGATGG 468
QY 593 ATATGCGCAGCTGGGGGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCCCTG 652
Db 467 ATATGCGCAGCTGGGGGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCCCTG 408
QY 653 GCTACGCGCAGCGGGGCGCATGAGAGGGGCGCGCCCTGACAGCAGGTCATCGGCC 712
Db 407 GCTACGCGCAGCGGGGCGCATGAGAGGGGCGCGCCCTGACAGCAGGTCATCGGCC 348
QY 713 ACTACCGGGGCTCTCTCTTCCAGACACGAGCAGTGGCGCCCTCTCTGCTGAGAG 772
Db 347 ACTACCGGGGCTCTCTCTTCCAGACACGAGCAGTGGCGCCCTCTCTGCTGAGAG 288
QY 773 GG 774
Db 287 GG 286
```

## RESULT 7

```
US-09-783-590-3464
; Sequence 3464, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruden, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
```

CURRENT APPLICATION NUMBER: US/09/783,590

CURRENT FILING DATE: 2000-02-15

PRIOR APPLICATION NUMBER: 08/420,856

PRIOR FILING DATE: 1995-04-12

PRIOR APPLICATION NUMBER: 08/346,731

PRIOR FILING DATE: 1994-11-21

NUMBER OF SEQ ID NOS: 12485

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3464

LENGTH: 368

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (103)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (225)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (279)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (314)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (349)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (350)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (366)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (367)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (368)

OTHER INFORMATION: n equals a,t,g, or c

US-09-783-590-3464

Query Match 11.8%; Score 102; DB 10; Length 368;

Best Local Similarity 100.0%; Pred. No. 1.9e-39; Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 715 TACCGGGGCTCTCTTCCAGACACGAGCAGTGGCGCCCTCTCTGAGAGGG 774
Db 1 TACCGGGGCTCTCTTCCAGACACGAGCAGTGGCGCCCTCTCTGAGAGGG 60
QY 775 ACCCGGCTCCACACACACATGCGCGCCCTAGAGAGCGCA 816
Db 61 ACCCGGCTCCACACACACATGCGCGCCCTAGAGAGCGCA 102
```

## RESULT 8

```
US-09-783-590-3488
; Sequence 3488, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruden, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
```

APPLICANT: Lee, Richard T.  
APPLICANT: Landschulz, Katherine T

```

RESULT 12
US-09-864-761-20542
; Sequence 20542, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US-09/864,761
; CURRENT FILING DATE: 2001-05-23
;
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21

```

RESULT 13  
US-09-864-761-3776  
Sequence 3776, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecomla-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIORITY APPLICATION NUMBER: US 60/180,312  
PRIORITY FILING DATE: 2000-02-04  
PRIORITY APPLICATION NUMBER: US 60/207,456  
PRIORITY FILING DATE: 2000-05-26  
PRIORITY APPLICATION NUMBER: US 09/632,366  
PRIORITY FILING DATE: 2000-08-03  
PRIORITY APPLICATION NUMBER: GB 24263.6  
PRIORITY FILING DATE: 2000-10-04  
PRIORITY APPLICATION NUMBER: US 60/236,359  
PRIORITY FILING DATE: 2000-09-27  
PRIORITY APPLICATION NUMBER: PCT/US01/00666  
PRIORITY FILING DATE: 2001-01-30  
PRIORITY APPLICATION NUMBER: PCT/US01/00667  
PRIORITY FILING DATE: 2001-01-30  
PRIORITY APPLICATION NUMBER: PCT/US01/00664  
PRIORITY FILING DATE: 2001-01-30  
PRIORITY APPLICATION NUMBER: PCT/US01/00669  
PRIORITY FILING DATE: 2001-01-30  
PRIORITY APPLICATION NUMBER: PCT/US01/00665  
PRIORITY FILING DATE: 2001-01-30  
PRIORITY APPLICATION NUMBER: PCT/US01/00668  
PRIORITY FILING DATE: 2001-01-30  
PRIORITY APPLICATION NUMBER: PCT/US01/00663  
PRIORITY FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3776
; LENGTH: 1964
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005291.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.7
; US-09-864-761-3776
```

```

Query Match      3.7%; Score 32; DB 10; Length 1964;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY 142 GTGGTGTGATGATGATGATGATGATGATGAT 173
DB 402 GTGGTGTGATGATGATGATGATGATGATGAT 433
```

```

RESULT 14
US-09-864-761-3936
; Sequence 3936, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: 2001-05-23
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3936
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007249.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.7
; US-09-864-761-3936
```

```

Query Match      3.4%; Score 29; DB 10; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY 143 TGGTGTGATGATGATGATGATGATGATGAT 171
DB 264 TGGTGTGATGATGATGATGATGATGATGAT 292
```

```

RESULT 15
US-09-864-761-20699
; Sequence 20699, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: 2001-05-23
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
```

```
? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: US 60/234,687
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 09/608,408
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: US 09/774,203
? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
? SEQ ID NO 20699
? LENGTH: 446
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AC007249.2
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 10
? OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 14
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 13
? OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
? OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12
? OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.7
? OTHER INFORMATION: NT HIT: AL161539.2, EVALU 3.70e+00
US-09-864-761-20699

Query Match 3.48; Score 29; DB 10; Length 446;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 143 TGGTGGTGGATGGTGGTGGTGGTGC 171
Db 143 TGGTGGTGGATGGTGGTGGTGGTGC 171
```

Search completed: March 18, 2003, 08:17:36  
Job time : 82 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 07:15:25 ; Search time 1293 Seconds  
(without alignments)  
10784.466 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861  
Sequence: 1 atgcacgcctgtgtggtggtt.....aacagaagaagacacctctc 861

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estnu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_inv: \*  
20: em\_gss\_pln: \*  
21: em\_gss\_vit: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_fod: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	614	71.3	967	14	B0641849
2	566	65.7	1046	14	B0622276
3	538	64.8	609	14	B0636742
4	550	63.9	551	13	B0141979
5	499	58.0	916	14	B0954555
6	497	57.7	890	14	B0690750

c	7	482	56.0	805	9	AL558881
c	8	470	54.6	729	14	B0575741
c	9	424	49.2	949	9	AL517150
c	10	420	48.8	655	14	B0691705
c	11	420	48.8	1280	14	B0691500
c	12	417	48.4	782	14	B0015170
c	13	417	48.4	898	9	AL558882
c	14	393	45.6	461	14	B0712680
c	15	389	45.2	633	14	B0714472
c	16	374	43.4	844	14	B0686793
c	17	369	42.9	559	12	BE855409
c	18	353	41.0	730	14	B0677602
c	19	351	40.8	626	14	B0974296
c	20	335	38.9	570	14	B0575582
c	21	330	38.3	552	14	B0713900
c	22	302	35.1	693	9	AL761441
c	23	279	32.4	547	14	B0676516
c	24	276	32.1	613	12	B0680325
c	25	257	29.8	308	14	B0705514
c	26	243	28.2	502	9	AL921394
c	27	238	27.6	446	14	B0681946
c	28	237	27.5	453	9	AA917446
c	29	235	27.3	437	9	AL936228
c	30	217	25.2	464	12	BF026695
c	31	206	23.9	518	9	AL885001
c	32	198	23.0	451	9	AL493698
c	33	198	23.0	990	12	B0675643
c	34	195	22.6	404	9	AL925027
c	35	194	22.5	634	9	AL826012
c	36	191	22.2	588	9	AL377498
c	37	191	22.2	619	9	AL742327
c	38	187	21.7	380	10	BE138909
c	39	183	21.3	508	9	AA595115
c	40	182	21.1	284	10	AA452945
c	41	179	20.8	467	10	AA007283
c	42	177	20.6	397	12	BF446904
c	43	177	20.6	430	10	AA204238
c	44	177	20.6	463	12	BF939262
c	45	177	20.6	674	9	AL972096

## ALIGNMENTS

RESULT 1  
B0641849  
LOCUS  
DEFINITION B0641849 967 bp mRNA linear EST 15-JUL-2002  
ACCESSION AGNCOURT\_8287174 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:6292265  
VERSION B0641849  
KEYWORDS B0641849.1 GI:21766021  
SOURCE EST.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 967)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strusberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)

FEATURES  
source  
1..967  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L10CM2493 row: 9 column: 18  
High quality sequence stop: 571.  
Location/Qualifiers



```

|||||
Db 427 GTGCGCCCTTGGCCAGGGGAGCGGCTTCACGGCTTCAGGCCACCTTCCGTACTG 486
QY 421 CAGCAGAGATGACCTGCCACCCACCATCTGCTGTGACAGCGGGAGAGCCCCACCC 480
Db 487 CACACAGAGATGACCTGCCAGCCACCACATCTGCTGTGACAGCGGGAGAGCCCCACCC 546
QY 481 TACACAGAGCCCTGCACCCCTGCAGCTTGGGAGCCCGAGAGAGAGTGAACGACCGG 540
Db 547 TACACAGAGCCCTGCACCCCTGCAGCTTGGGAGCCCGAGAGAGAGTGAACGACCGG 606
QY 541 GACTCGGTGGCGACCCCAACAGAACATCTGACAGATGACCTGATGATAGTGC 600
Db 607 GATCGGTGGCGACCCCAACAGAACATCTTGCAGATGACCTGATGATAGTGC 666
QY 601 AGGCTGGGGGGCCCTGCCCCCAGACGATGCTGGGGCATCAGCCGACGCTCTACGC 660
Db 667 AGGCTGGGGGGCCCTGCCCCCAGACGATGCTGGGGCATCAGCCGACGCTCTACGC 726
QY 661 AGGGGGG 668
Db 727 AGCGCGG 734

RESULT 3
B0636742 609 bp mRNA linear EST 15-JUL-2002
LOCUS hd13h06.y1 Human Retina cDNA (un-normalized, unamplified): hd/he
DEFINITION Homo sapiens cDNA clone hd13h06 5', mRNA sequence.
ACCESSION B0636742
VERSION B0636742.1 GI:21761201
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 609)
AUTHORS Wislow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behl,A., Touchman
J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
JOURNAL Contact: Wislow G
COMMENT Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 13 row: h column: 06
Seq primer: M13Rpl reverse primer (ABI).
FEATURES
source
1. 609
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hd13h06"
/clone_lib="Human Retina cDNA (un-normalized, unamplified)
): hd/he"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDHL08"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the Superscript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-pGACTAGTCTAGATCGGAGCGGCGCC(T)15-3'

```

```

BASE COUNT 114 a 238 c 182 g 75 t
ORIGIN
Query Match 64.8%; Score 558; DB 14; Length 609;
Best local Similarity 99.8%; Pred. No. 3.1e-234;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 219 CAGCCGGACAGCCAGGGGGGAGAGAGAAGATGCCCTGCTCAGAAAGATGCCGTG 278
Db 1 CAGCCGGACAGCCAGGGGGGAGAGAGAAGATGCCCTGCTCAGAAAGATGCCGTG 60
QY 279 GCCCTGGAGAGACAGATGTGACGCAACGGAATGCCAGAGCCGAGTCTACGCCGCC 338
Db 61 GCCCTGGAGAGACAGATGTGACGCAACGGAATGCCAGAGCCGAGTCTACGCCGCC 120
QY 339 TCGGCCACGACCGCCCTGAGCCGTCGCCCTTGCCACGAGGAGGCGTTCACGCGCTT 398
Db 121 TCGGCCACGACCGCCCTGAGCCGTCGCCCTTGCCACGAGGAGGCGTTCACGCGCTT 180
QY 399 CCAAGCCACCTATCCGTACCTGACAGACAGATGACCTGCAACCCACATCTGCTGTC 458
Db 181 CCAAGCCACCTATCCGTACCTGACAGACAGATGACCTGCGGCCACATCTGCTGTC 240
QY 459 AGACGGGGAGAGAGCCCCACACCTACAGAGGCCCTGACCTTCAGTTCGGAGCCGA 518
Db 241 AGACGGGGAGAGAGCCCCACACCTACAGAGGCCCTGACCTTCAGTTCGGAGCCGA 300
QY 519 GCAGACGCTGGAAGTGAACCGGGGAGTGGTGGCGGCACCCCAACAGAACATCTTGA 578
Db 301 GCAGACGCTGGAAGTGAACCGGGGAGTGGTGGCGGCACCCCAACAGAACATCTTGA 360
QY 579 CAGTGACCTGATGATAGTATGATGACAGGCTGGGCGGCCCTGCCCCCAGCAGTAATCGG 638
Db 361 CAGTGACCTGATGATAGTATGATGACAGGCTGGGCGGCCCTGCCCCCAGCAGTAATCGG 420
QY 639 CATTACGGCCACATGCTGCTACGCGAGAGCGGGCGGCATGAGAGGGCGCGCCACCTACG 698
Db 421 CATTACGGCCACATGCTGCTACGCGAGCGGGCGGCATGAGAGGGCGCGCCACCTACG 480
QY 699 CGAGGTCATCGGCGACACTACCGGGGCTCCTCTCCACAGCAGCAGAGAGAGTGGCGGCC 758
Db 481 CGAGGTCATCGGCGACACTACCGGGGCTCCTCTCCACAGCAGCAGAGAGAGTGGCGGCC 540
QY 759 CTCCTTGTGAGAGGGGACCCGGCTCCACACACACATGCGCGCCCTAGAGAGCGACG 818
Db 541 CTCCTTGTGAGAGGGGACCCGGCTCCACACACACATGCGCGCCCTAGAGAGCGACG 600
QY 819 CATCTGGAG 827
Db 601 CATCTGGAG 609

RESULT 4
BM141979 551 bp mRNA linear EST 12-MAR-2002
LOCUS BM141979
DEFINITION I25a11.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5673341 5' similar to TR:09UDJ3 09UDJ3 DJ718J7.1
; , mRNA sequence.
ACCESSION BM141979
VERSION BM141979.1 GI:117152046
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 551)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Seacrest,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistein,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagarisvilli,R., Williams,T.

```

TITLE  
Journal  
COMMENT

, Jackson, Y. and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biolhp.harvard.edu  
Library was constructed by Dr. Douglas Melton DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Juliana Brown  
(brownjefas.harvard.edu) This sequence now available from the IMAGE  
consortium, for clone orders contact: info@image.llnl.gov  
Seq primer: -40RP from Glibco  
High quality sequence stop: 429.

## FEATURES

source

1..551  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5677341"  
/clone\_lib="Melton Normalized Human Islet 4 N4-HIS 1"  
/sex="Both"  
/tissue\_type="Islets of Langerhans"  
/dev\_stage="Adult"  
/lab\_host="DH10B"  
/note="Organ: Pancreas; Vector: pSPORT1; Site\_1: Not 1;  
Site\_2: Sal 1; Starting library constructed using  
SuperScript Plasmid Library kit (Life Technologies). cDNA  
made by oligo-dT priming. Size-selected by column  
fractionation; average insert size 1.08 kb. Library was  
amplified once on solid support and plasmid DNA from  
library was prepared. The library DNA was normalized by  
method #4 from Bonaldo, Lennon, and Soares 1996 Genome  
Research 6:791-806; 0.5 microgram single-stranded library  
plasmid DNA was mixed with 5 micrograms PCR product  
representing library inserts and hybridized to an EcoT of  
20. Single-stranded (unhybridized) plasmids were isolated  
by hydroxyapatite chromatography and used to make this  
library."

BASE COUNT 110 a 218 c 157 g 66 t

ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 550; DB 13; Length 551;  
Matches 550; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

302 GCACGGAATCCAGAGCGGAGTCTACGCGCCGCTGCGCCACCGACGCGCTGGCGG 361  
1 GCACGGAATCCAGAGCGGAGTCTACGCGCCGCTGCGCCACCGACGCGCTGGCGG 60  
302 TGCCGCGCTTGGCCAGCGGAGGCGCTTCACCGCTTCAGCCACCTATTCGTAACCTGC 421  
61 TGCCGCGCTTGGCCAGCGGAGGCGCTTCACCGCTTCAGCCACCTATTCGTAACCTGC 120  
422 AGCAGAGATGAGCTGCGACCGACCATCTGCTGTCAGACGGGAGAGAGAGAGAGAGAGAG 481  
121 AGCAGAGATGAGCTGCGACCGACCATCTGCTGTCAGACGGGAGAGAGAGAGAGAGAGAG 180  
482 ACCAGAGGCGCTGACCGCTTCAGCTTGGGAGACCGCCGAGCAGAGCTGGAATGAAACCGGG 541  
181 ACCAGAGGCGCTGACCGCTTCAGCTTGGGAGACCGCCGAGCAGAGCTGGAATGAAACCGGG 240  
542 AGTCGCTGCGGCGACCGCCAAACAGACCATCTTGCAGACGATGATGATGATGATGATGCA 601  
241 AGTCGCTGCGGCGACCGCCAAACAGACCATCTTGCAGACGATGATGATGATGATGATGCA 300  
602 GCGTGGGCGGCG 661  
301 GCGTGGGCGGCG 360

QY 662 GCGGCGGCGCATGAGAGGCGCGCGCCACCTACAGCGGAGTCACTGGGCGCACTACCGG 721  
DB 361 GCGGCGGCGCATGAGAGGCGCGCGCCACCTACAGCGGAGTCACTGGGCGCACTACCGG 420  
QY 722 GGTCTCTCTTCAGCAGCAGCAGAGCAGTGGGCGCGCTCTCTCTCTCTCTCTCTCTCTCT 781  
DB 421 GGTCTCTCTTCAGCAGCAGCAGAGCAGTGGGCGCGCTCTCTCTCTCTCTCTCTCTCTCT 480  
QY 782 TCACACACACACATCGCGCGCGCTTACAGAGCGGACCGCATCTGAGCAAGAGAGATGA 841  
DB 481 TCACACACACACATCGCGCGCGCTTACAGAGCGGACCGCATCTGAGCAAGAGAGATGA 540  
QY 842 AACAGAAAG 851  
DB 541 AACAGAAAG 550

## RESULT 5

B0954555

## LOCUS

B0954555 916 bp mRNA linear EST 21-AUG-2002

## DEFINITION

AGENCOURT\_8825282 lupskl\_sciatic\_nerve Homo sapiens cDNA clone

## ACCESSION

B0954555  
IMAGE:6204609 5', mRNA sequence.

## VERSION

B0954555.1 GI:22370033

## KEYWORDS

EST.

## SOURCE

human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

## REFERENCE

1 (bases 1 to 916)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

## AUTHORS

Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

## TITLE

cDNA Library Preparation: Life Technologies, Inc.

## JOURNAL

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

## COMMENT

Plate: LLAM13626 row: c column: 10  
High quality sequence stop: 669.

## FEATURES

source

Location/Qualifiers  
1..916  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6204609"  
/clone\_lib="lupskl\_sciatic\_nerve"  
/sex="male"  
/tissue\_type="sciatic nerve"  
/dev\_stage="adult, 70 yr"  
/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
Not1; Site\_2: Sal1; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TGACACGAGCGTCG-3' and  
5'-GACTAGTCTAGATCGGAGCGGCGCGCTT(15)-3'. Size selected >  
1 kb for average insert length 1.87 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies."

## BASE COUNT

162 a 354 c 273 g 127 t

## ORIGIN

## Query Match

Best Local Similarity 99.7%; Score 499; DB 14; Length 916;  
Matches 599; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGACCGCTTGTATGGGGGTCAACAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60  
|||||

[illegible]

```

/adb_xref="taxon.9606"
/clone="IMAGE:6209341"
/clone_id="NIH_MGC_110"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT      168 a      303 c      273 g      146 t
ORIGIN

Query Match      57.7%; Score 497; DB 14; Length 890;
Best Local Similarity 99.8%; Pred. No. 1.8e-207;
Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  110  CGGAGCTGGAGTTGTTCATCATCATCATCGTGGTGTATGTATGTGTGTG 169
      |||
Db   128  CGGAGCTGAGTTGTTCATCATCATCATCGTGGTGTATGTATGTGTGTG 187

QY  170  TGTATCACTGCTGCTGTGAGCCACTCAAGCTGTCTGACAGGTCCTTCAATCAGCCGGCACA 229
      |||
Db   188  TGTATCACTGCTGCTGTGAGCCACTCAAGCTGTCTGACAGGTCCTTCAATCAGCCGGCACA 247

QY  230  GCCAGGGGCGGAGGAGAAAGATGCCCTGCTCGAAGAGATGCTGTGGCCCTCGGAGAGA 289
      |||
Db   248  GCCAGGGGCGGAGGAGAAAGATGCCCTGCTCGAAGAGATGCTGTGGCCCTCGGAGAGA 307

QY  290  GCACAGTGTCAAGCAACGGAATCCAGAGCCGCAGGTCTACGCCGCCGCTCGGCCACCG 349
      |||
Db   308  GCACAGTGTCAAGCAACGGAATCCAGAGCCGCAGGTCTACGCCGCCGCTCGGCCACCG 367

QY  350  ACCGCTGTGGCGGTGCGCCCTTCGCCACAGGGGAGCGCTTCACAGCGCTTCACAGCCCACT 409
      |||
Db   368  ACCGCTGTGGCGGTGCGCCCTTCGCCACAGGGGAGCGCTTCACAGCGCTTCACAGCCCACT 427

QY  410  ATCCGTACTCCTGAGACAGATGCATGCACTGCCACCCACCACATCTCGCTGTTCAGACGGGAGG 469
      |||
Db   428  ATCCGTACTCCTGAGACAGATGCATGCACTGCCGCCCCACCATCTCGCTGTTCAGACGGGAGG 487

QY  470  AGCCCCCAACCTTACAGAGGGCCCTCGACACCTTCACAGCTTCGGGAGCCCGAGACAGCACTGG 529
      |||
Db   488  AGCCCCCAACCTTACAGAGGGCCCTCGACACCTTCACAGCTTCGGGAGCCCGAGACAGCACTGG 547

QY  530  AACTGTAACCGGGAGTGTGCTGCGCGCACCCCAACAAGACAATCTTTCGACAGTCACTGA 589
      |||
Db   548  AACTGTAACCGGGAGTGTGCTGCGCGCACCCCAACAAGACAATCTTTCGACAGTCACTGA 607

QY  590  TGGATAGTGCACAGGTGTGGGGGCCCTCGCCGCCCCAGACAGTAACTCGGGGCATCAACGGCA 649
      |||
Db   608  TGGATAGTGCACAGGTGTGGGGGCCCTCGCCGCCCCAGACAGTAACTCGGGGCATCAACGGCA 667

QY  650  CGTGTCTAC 657
      |||
Db   668  CGTGTCTAC 675

RESULT 7
AL558881/c 805 bp mRNA linear EST 16-FEB-2001
LOCUS LOCUS
DEFINITION AL558881.L1.L_NFL008.TC2 Homo sapiens cDNA clone G50D015F12 3
ACCESSION AL558881
VERSION AL558881.1 GI:12903836
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE 1 (bases 1 to 805)  
 AUTHORS Li W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES  
 source location/Qualifiers

1..805

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CSDBJ015YF12"

/clone\_1lb="UTL-NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax: (1) 301 610 8371

Email: fliang@lifetech.com URL:

http://fulllength.invitrogen.com"

BASE COUNT 106 a 248 c 289 g 161 t 1 others

ORIGIN

Query Match 56.0%; Score 482; DB 9; Length 805;

Best Local Similarity 99.5%; Pred. No. 6.9e-201;

Matches 632; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

227 ACAGCCAGGCGGAGAGAGAGATGCCCTGCTCAGAGAGATCCCTGTCGG 286  
 DB ACAGCCAGGCGGAGAGAGAGATGCCCTGCTCAGAGAGATCCCTGTCGG 696  
 287 AGAGCAGATGTCAGGCAAGGAAATCCAGAGCGGAGGCTTACGCCGCTCGGCCA 346  
 DB AGAGCAGATGTCAGGCAAGGAAATCCAGAGCGGAGGCTTACGCCGCTCGGCCA 636  
 695 AGAGCAGATGTCAGGCAAGGAAATCCAGAGCGGAGGCTTACGCCGCTCGGCCA 636  
 347 CGAGCCGCTGCGCGCTGCGCCCTTCCAGCGGAGGCGCTTCCAGCGCTTCCAGCCCA 406  
 DB CGAGCCGCTGCGCGCTGCGCCCTTCCAGCGGAGGCGCTTCCAGCGCTTCCAGCCCA 576  
 635 CGAGCCGCTGCGCGCTGCGCCCTTCCAGCGGAGGCGCTTCCAGCGCTTCCAGCCCA 576  
 407 CCTATCCGATCTGACAGCAGAGATGAGCTGCAACCCATCTCTGCTTACAGAGGGG 466  
 DB CCTATCCGATCTGACAGCAGAGATGAGCTGCAACCCATCTCTGCTTACAGAGGGG 516  
 575 CCTATCCGATCTGACAGCAGAGATGAGCTGCAACCCATCTCTGCTTACAGAGGGG 516  
 467 AGAGCCGCTGCGCGCTGCGCCCTTCCAGCGGAGGCGCTTCCAGCGCTTCCAGCCCA 526  
 DB AGAGCCGCTGCGCGCTGCGCCCTTCCAGCGGAGGCGCTTCCAGCGCTTCCAGCCCA 456  
 515 AGAGCCGCTGCGCGCTGCGCCCTTCCAGCGGAGGCGCTTCCAGCGCTTCCAGCCCA 456  
 527 TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 586  
 DB TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 396  
 455 TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 396  
 587 TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 646  
 DB TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 336  
 395 TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 336  
 647 CCAGTGTACGAGCAGCGGCGGCGATGAGAGGCGCGCCGCTTCCAGCGGAGTGAAC 706  
 DB CCAGTGTACGAGCAGCGGCGGCGGCGATGAGAGGCGCGCCGCTTCCAGCGGAGTGAAC 276  
 335 CCAGTGTACGAGCAGCGGCGGCGGCGATGAGAGGCGCGCCGCTTCCAGCGGAGTGAAC 276  
 DB CCAGTGTACGAGCAGCGGCGGCGGCGGCGATGAGAGGCGCGCCGCTTCCAGCGGAGTGAAC 216  
 275 TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 216  
 707 TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 766  
 DB TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 826  
 767 TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 826  
 DB TGGAGTGAACCGGAGGAGTGGTGGCGCACCCCAACAGAACATCTTGCAGAGTGAAC 156

QY 827 GCAAGAGAGATTAACAGAGACACCTCTC 861  
 DB 155 GCAAGAGAGATTAACAGAGACACCTCTC 121

RESULT 8

BO575741/c

LOCUS

DEFINITION

BO575741

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

FEATURES  
 source location/Qualifiers

1..729

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-H-E21-bdg-h-14-0-UI"

/clone\_1lb="NCI-CGAP\_Ch2"

/tissue\_type="Chondrosarcoma Grade II"

/dev\_stage="Adult"

/lab\_host="DHI0B (Life Technologies)"

/note="Organ: Left pelvis; Vector: p773-Pac (Pharmacia)

with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I;

NCI-CGAP\_Ch2 is a normalized cDNA library containing the

following tissue(s): Chondrosarcoma Grade II. The library

was constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into p773-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

(dT)18 tail. The sequence tag for this library is

TGATCAGCCT.

TAG\_LIB-UI-H-E21

TAG\_TISSUE-grade-2-chondrosarcoma

TAG\_SEQ-ATCTAATATG"

BASE COUNT 101 a 213 c 253 g 162 t

ORIGIN

Query Match 54.6%; Score 470; DB 14; Length 729;

Best Local Similarity 99.8%; Pred. No. 1.3e-195;

Matches 520; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 341 GGCCACGACGCGCTGCGCGCTTCCAGCGGAGGCGCTTCCAGCGCTTCC 400  
 DB 729 GGCCACGACGCGCTGCGCGCTTCCAGCGGAGGCGCTTCCAGCGCTTCC 670  
 QY 401 AGCCACCTATCCGATCTGACAGCAGAGATGAGCTGCGCACCCACACATCTGCTCAG 460  
 DB 669 AGCCACCTATCCGATCTGACAGCAGAGATGAGCTGCGCACCCACACATCTGCTCAG 610



cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov

Plate: L1CM2367 row: 9 column: 20  
High quality sequence stop: 645.

## FEATURES

Location/Qualifiers

1..655  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6208939"  
/clone\_lib="NIH\_MGC\_110"  
/tissue\_type="ductal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

BASE COUNT 155 a 216 c 193 g 91 t

## ORIGIN

Query Match 48.8%; Score 420; DB 14; Length 655;  
Best Local Similarity 100.0%; Pred. No. 1.1e-173;

Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

442 CCCACATCTGCTGTCAAGCGGAGAGAGCCGCCCTACAGAGGCCCTGACACCTC 501  
28 CCCACATCTGCTGTCAAGCGGAGAGAGCCGCCCTACAGAGGCCCTGACACCTC 87  
502 CAGCTTGGGAGCCCGAGAGAGAGTGAACCGGAGTGGTGGCGCCACCCCA 561  
88 CAGCTTGGGAGCCCGAGAGAGAGTGAACCGGAGTGGTGGCGCCACCCCA 147  
552 AACAGAACATCTTGCAGAGTGAATGATGATGATGATGATGATGATGATGAT 621  
148 AACAGAACATCTTGCAGAGTGAATGATGATGATGATGATGATGATGATGAT 207  
622 CCCACATCTGCTGTCAAGCGGAGAGAGCCGCCCTACAGAGGCCCTGACACCTC 681  
208 CCCACATCTGCTGTCAAGCGGAGAGAGCCGCCCTACAGAGGCCCTGACACCTC 267  
682 CCGCGCCGCCCTACAGAGAGAGTCAATGAGGAGTCAATGAGGAGTCAATGAG 741  
268 CCGCGCCGCCCTACAGAGAGAGTCAATGAGGAGTCAATGAGGAGTCAATGAG 327  
742 CAGAGCACTGGGCGCCCTCTGCTGTGAGAGGAGCCCGGCTCCACACACATCGCG 801  
328 CAGAGCACTGGGCGCCCTCTGCTGTGAGAGGAGCCCGGCTCCACACACATCGCG 387  
802 CCCCTAGAGAGCGAGCCATCTGAGAGCAAAAGAGATTAACAGAAAGGACACCTCTC 861  
388 CCCCTAGAGAGCGAGCCATCTGAGAGCAAAAGAGATTAACAGAAAGGACACCTCTC 447

RESULT 11  
LOCUS B0691500 1280 bp mRNA linear EST 15-JUL-2002  
DEFINITION AGNCOURT\_8345477 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6248217  
5', mRNA sequence.  
ACCESSION B0691500  
VERSION B0691500.1 GI:21816816  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1280)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov

Plate: L1CM2367 row: 1 column: 10  
High quality sequence stop: 423.

## FEATURES

Location/Qualifiers

1..1280  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6248217"  
/clone\_lib="NIH\_MGC\_110"  
/tissue\_type="ductal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

BASE COUNT 298 a 423 c 374 g 182 t 3 others

## ORIGIN

Query Match 48.8%; Score 420; DB 14; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 1.1e-173;

Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

442 CCCACATCTGCTGTCAAGCGGAGAGAGCCGCCCTACAGAGGCCCTGACACCTC 501  
28 CCCACATCTGCTGTCAAGCGGAGAGAGCCGCCCTACAGAGGCCCTGACACCTC 87  
502 CAGCTTGGGAGCCCGAGAGAGAGTGAACCGGAGTGGTGGCGCCACCCCA 561  
88 CAGCTTGGGAGCCCGAGAGAGAGTGAACCGGAGTGGTGGCGCCACCCCA 147  
552 AACAGAACATCTTGCAGAGTGAATGATGATGATGATGATGATGATGATGAT 621  
148 AACAGAACATCTTGCAGAGTGAATGATGATGATGATGATGATGATGATGAT 207  
622 CCCACATCTGCTGTCAAGCGGAGAGAGCCGCCCTACAGAGGCCCTGACACCTC 681  
208 CCCACATCTGCTGTCAAGCGGAGAGAGCCGCCCTACAGAGGCCCTGACACCTC 267  
682 CCGCGCCGCCCTACAGAGAGAGTCAATGAGGAGTCAATGAGGAGTCAATGAG 741  
268 CCGCGCCGCCCTACAGAGAGAGTCAATGAGGAGTCAATGAGGAGTCAATGAG 327  
742 CAGAGCACTGGGCGCCCTCTGCTGTGAGAGGAGCCCGGCTCCACACACATCGCG 801  
328 CAGAGCACTGGGCGCCCTCTGCTGTGAGAGGAGCCCGGCTCCACACACATCGCG 387  
802 CCCCTAGAGAGCGAGCCATCTGAGAGCAAAAGAGATTAACAGAAAGGACACCTCTC 861  
388 CCCCTAGAGAGCGAGCCATCTGAGAGCAAAAGAGATTAACAGAAAGGACACCTCTC 447

RESULT 12  
LOCUS B0015170 782 bp mRNA linear EST 26-MAR-2002  
DEFINITION UI-H-ED1-axw-k-20-0-UI.s1 NCI CGAP\_Ed1 Homo sapiens cDNA clone  
IMAGE:5834655 3', mRNA sequence.  
ACCESSION B0015170  
VERSION B0015170.1 GI:19740071  
KEYWORDS EST.



SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Homo sapiens: Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
COMMENT 1 (bases 1 to 782)  
Tumor Gene Index  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapds-femail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found  
through the I.M.A.G.E. Consortium/LINL at: http://image.lnl.gov  
Seq primer: M13 FORWARD  
POLYA-Yes.

FEATURES  
source Location/Qualifiers  
1..782  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5834635"  
/clone\_1id="NCI\_CGAP\_ED1"  
/tissue\_type="Chondrosarcoma"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Organ: Left Pubic Bone; Vector: pT73-Pac  
(Pharmacia) with a modified polylinker; Site 1: EcoR I;  
Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library  
containing the following tissue(s): Chondrosarcoma cell  
line C5. The library was constructed according to Bonaldo  
, Lennon and Soares, Genome Research, 6:791-806, 1996.  
First strand cDNA synthesis was primed with an oligo-dT  
primer containing a Not I site. Double stranded cDNA was  
ligated to an EcoR I adaptor, digested with Not I, and  
cloned directionally into pT73-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is GCTCAAGCTT.  
TAG LIB=UI-H-ED1  
TAG TISSUE=chondrosarcoma  
TAG\_SEQ=CCTCAAGCTT"

BASE COUNT 109 a 223 c 271 g 176 t 3 others  
ORIGIN

Query Match 48.4%; Score 417; DB 14; Length 782;  
Best Local Similarity 99.6%; Pred. No. 2.3e-172;  
Matches 517; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 343 CCCACGACCGCTGGCGCTGCGCCCTTCCGAGGGAGCGCTTCCACCGCTTCAG 402  
|||||  
Db 725 CCCACGACCGCTGGCGCTGCGCCCTTCCGAGGGAGCGCTTCCACCGCTTCAG 666  
|||||

QY 403 CCCACCTATCCGTACTGACAGAGATGCACTGCGCACCACCATCTGCTGACAG 462  
|||||  
Db 665 CCCACCTATCCGTACTGACAGAGATGCACTGCGCACCACCATCTGCTGACAG 606  
|||||

QY 463 GGGGAGAGAGCCCACTTACAGAGGCGCTGACACCTTCAGCTTGGAGACCCGAGAG 522  
|||||  
Db 605 GGNAGAGAGAGCCCACTTACAGAGGCGCTGACACCTTCAGCTTGGAGACCCGAGAG 546  
|||||

QY 523 CAGCTGAACTGAACCGGAGTGGTGGCGGACCCCAACAGAACCATCTTGCACAGT 582  
|||||  
Db 545 CAGCTGAACTGAACCGGAGTGGTGGCGGACCCCAACAGAACCATCTTGCACAGT 486  
|||||

QY 583 GACCTGATGATGATGCGAGGCTGGGGGCGCCCGCCCGCCAGAGAAATCGGGCATC 642  
|||||  
Db 485 GACCTGATGATGATGCGAGGCTGGGGGCGCCCGCCCGCCAGAGAAATCGGGCATC 426  
|||||

QY 643 AGCCGACGTGCTAGCGGAGGGGCGCATGAGAGGGGCGCGCCACCTACAGCGAG 702  
|||||

Db 425 AGCCGACGTGCTAGCGGAGGGGCGCATGAGAGGGGCGCGCCACCTACAGCGAG 366  
|||||

QY 703 GTCATCGGCACATACCCGGGCTCTCTTCCAGACACCGAGAGAGATGGGCCCTTCC 762  
|||||

Db 365 GTCATCGGCACATACCCGGGCTCTCTTCCAGACACCGAGAGAGATGGGCCCTTCC 306  
|||||

QY 763 TTGCTGGAGGGGACCGGCTCCACACACATCGGCGCCCTGAGAGGGCAGCCATC 822  
|||||

Db 305 TTGCTGGAGGGGACCGGCTCCACACACATCGGCGCCCTGAGAGGGCAGCCATC 246  
|||||

QY 823 TGGAGCAAGAGAGATTAACAGAAAGAGACCCCTTC 861  
|||||

Db 245 TGGAGCAAGAGAGATTAACAGAAAGAGACCCCTTC 207  
|||||

RESULT 13  
AL558882 898 bp mRNA linear EST 16-FEB-2001  
LOCUS AL558882 LTI\_NFL008.TC2 Homo sapiens cDNA clone CS0D015YF12 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL558882  
VERSION AL558882.1 GI:12903838  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1. (bases 1 to 898)  
AUTHORS Li, W.-B., Gruber, C., Jesssee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source Location/Qualifiers  
1..898  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D015YF12"  
/clone\_1id="LTI\_NFL008\_TC2"  
/sex="male"  
/tissue\_type="T cells from T cell leukemia"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
vector. Library was normalized. Library was constructed by  
Life Technologies. Contact : Feng Liang Life Technologies,  
a division of Invitrogen 9800 Medical Center Drive  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 196 a 310 c 260 g 126 t 6 others  
ORIGIN

Query Match 48.4%; Score 417; DB 9; Length 898;  
Best Local Similarity 99.5%; Pred. No. 2.3e-172;  
Matches 637; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 177 GTGCTGCTGAGCCACTCAAGCTGTGCAACGGCTTCATCAACCGCGCAGCAGAGG 236  
|||||

Db 1 GTGCTGCTGAGCCACTCAAGCTGTGCAACGGCTTCATCAACCGCGCAGCAGAGG 60  
|||||

QY 237 GCGAGGAGAGAGATGCTGCTCTCAGAGGATGCTGTGGCCCTGGAGAGACAGT 296  
|||||

Db 61 GCGAGGAGAGAGATGCTGCTCTCAGAGGATGCTGTGGCCCTGGAGAGACAGT 120  
|||||

QY 297 GTCAAGCAAGGAATCCAGAGCGCAGGTCTAGCGCCCGCCCTGGGCCACGACGCT 356  
|||||

Db 121 GTCAAGCAAGGAATCCAGAGCGCA-GTCTACGCCCCCGCTGGGCCACGACGCT 179  
|||||

QY	357	GGCGGTGGCGGCTTCTGGCGCCAGCGGAGCGCTTCCAGCGCTTCCAGCCACCTATTCGTA	416
Db	180	GGCCGTGCCGCCCTTCTGCCCCAGCGGAGCGCTTCCAGCGCTTCCAGCCACCTATTCGTA	239
QY	417	CCTGCAGCAGAGATGCAGCTTCGACCCACCATCTCGCTGTCCAGCGGGGAGAGCCCC	476
Db	240	CCTGCAGCAGCAGATCGACCTTCGCGGCCACCATCTCGCTGTCCAGCGGGGAGAGCCCC	299
QY	477	ACCCACACGAGGGCCCTGCACCTTCACCTTCGGGAGCCCGCAGCAGCAGTGGAACTGAA	536
Db	300	ACCCATACGAGGGCCCTGCACCTTCAGCTTCGGGAGCCCGCAGCAGCAGTGGAACTGAA	359
QY	537	CCGGGAGTCGGTGGCGCGCACCCGCCAAACAGAACCATCTTGCACAGTGCATGTATGATAG	596
Db	360	CCGGGAGTGGGTGGCGCGCACCCGCCAAACAGAACCATCTTGCACAGTGCATGTATGATAG	419
QY	597	TGCCAGGCTGGGGCGGGCCCTGCCCCCGCCAGCAGTAACTCGGGCATTACAGCGCAGCTGGTA	656
Db	420	TGCCAGGCTGGGGCGGGCCCTGCCCCCGCCAGCAGTAACTCGGGCATTACAGCGCAGCTGGTA	479
QY	657	CGCGCAGCGGCGGCGCATGAGAGGGGCGCGGCCACTACAGCGAGGTATCGCGCAGCTA	716
Db	480	CGCGCAGCGGCGGCGCATGAGAGGGGCGCGGCCACTACAGCGAGGTATCGCGCAGCTA	539
QY	717	CCCGGGGTCCTCTTCCAGCAGCAGCAGCAGCAGTGGGCGGCCCTCTCTTGTCTGGAGGGAC	776
Db	540	CCCGGGGTCCTCTTCCAGCAGCAGCAGCAGCAGTGGGCGGCCCTCTCTTGTGGAGGGAC	599
QY	777	CCGGCTCCACACACACACATCGGGGCCCTGTAGAGGCCA	816
Db	600	CCGGCTCCACACACACACATCGGGGCCCTGTAGAGGCCA	639

Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker. Site\_1: EcoR I; Site\_2: Not I;  
UR-E-EJ0 is a subtracted cDNA library constructed  
according to Bonaldi, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT7T3-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are: fetal eyes, AGAATCAAGA  
; lens, CGAATGACGA; eye anterior segment, AATGCCGCT;  
optic nerve, CCATTAAGTG; retina, CCGCG; Retina Foveal and  
Macular, GTCC; RPE and Choroid, ACCCTA. This library was  
created for the program, Gene Discovery in the Visual  
System supported by National Eye Institute (NEI)."

Query Match	45.68;	Score 393;	DB 14;	Length 461;
Best Local Similarity	100.0%;	Pred. No. 7.7e-162;		
Matches 393;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 469 GACCCCCCACCCTTACCAGGGCCCTGCACCCCTCCAGTTCGGGACCCCGAGCAGCTG 528

529 GAACTGAACCGGAGTGGTGGCGCGCACCCCCAACAAGAACCATCTTGACAGTACCTG 588

0Y 589 ATGGATAGTGGCAGGCTGGGGCGCCCTGCCCCCAGCAGTAACTCGGGCGATCAGCGCC 648  
 DB b1 GCACTGAAACGGGAGTGGGTGGCGGCAACCCCCAACAGAACCTATTCGACAGTAGCCCG 120

Db 121 ATGGATAGTGGCAGGCTGGGCGGCCCTGCCCCCAGCAGTAACTCGGGCAATCAAGCGCC 180

Db 181 ACGTGCTACGGAGCGGGCGGCATGAGAGGGCCGCGGCCACACTACAGGAGGTATC 240

QY 709 GGCCACTACCCGGGGTCCCTCTTCCAGCACACGACAGACAGTGGGCGGCCCTCCTTGCTG 768

769 GAGGGGACCCGGCTCCACCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGC 828

829 **AAAGGAAAGGATAAACAGAAAGGACACCCCTTC** 861  
Db 301 **GAGGGGACCCCGGCTCCACCACACACACATCGCGCCCTTAGAGAGCCGACCCATCTGGAGC** 350

Db 361 AAAGAGAAGGATTAACAGAGAAGACACCCCTCTC 393

RESULT 15  
BM714472

**DEFINITION.** UI-E-EJ0-ahs-f-20-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone UI-E-EJ0-ahs-f-20-0-UI 5', mRNA sequence.

```

REVISION 2011.11.25  GI:19027730
VERSION BM714472.1
KEYWORDS EST.

```

ORGANISM	Homio sapiens	Chordata: Vertebrata: Euteleostei	Chordata: Vertebrata: Homio sapiens
Eukaryota; Metazoa;			

REFERENCE  
1 (bases 1 to 633)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate a

JOURNAL of discovery  
Genome Res. 6 (9), 791-806 (1996)

Tue Mar 18 11:49:33 2003

us-09-934-249-3.oli.rst

Page 11

**MEDLINE  
COMMENT**

97044477  
Contact:

Contact: Soares, MB

# Program for Rat Gene Discovery and Mapping

University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel.: 319 335 8250

Fax: 319 335 9565

Fax: 319 335 9565

Fax: 319 335 9565

Fax: 319 335 9565

Email: [msoares@blue.weeg.iowa.edu](mailto:msoares@blue.weeg.iowa.edu)

CDNA Library Preparation: Dr. M. Ben  
I Issue Procurement: Dr. M. Ben  
I Issue Procurement: Dr. M. Ben

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Genetics ([www.resqgen.com](http://www.resqgen.com)): Researchers may obtain clones from ResqGen

The following repetitive elements were found in this cDNA

sequence: 593-614, >AT\_rich#Low\_complexity (matched compliment)

seq primer: M13 reverse.  
Location/Qualif

## FEATURES

**Source**

1. .633

/organism="Homo sapiens"

```
/db_xref="taxon:9606"
/cname="III-E-ETD-abs-
```

```
/clone_lib="UI-E-EJ0"
```

```
/tissue_type="fetal e
```

optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"

```
/dev_stage="fetal and adult"
```

```
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/notes="Organ: Yector: mTm3-Bac (Pharmacia) with a"
```

modified polylinker: site 1: EcoR I; site 2: Not I;  
/note-organ: eye, vector: pRIS-fac (Pharmacia) within a

BASE COUNT  
ORIGIN

150 a	206 c	182 g	93 t	2 others
-------	-------	-------	------	----------

## ORIGIN

Query match

45.28; Score 389; DB 14; Length 633;

Best Local Similarity 100.0%; Pred. NO. 4.4e-160;  
Matches 389: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy	473	CCCCACCCATTACAGAGGGCCCTTGACACCTCCAGCTCTTGGGAGCCCGAGACACACCTGGAAAC	53.2
Db	14	CCCCACCCATTACAGAGGGCCCTTGACACCTCCAGCTCTTGGGAGCCCGAGACACACCTGGAAAC	73
Oy	533	TGAACCGGAGTGTGGTGGCGCACCCCCCAACAGAAACATCTTGTGACAGTGAACCTGATGG	59.2
Db	74	TGAACCGGAGTGTGGTGGCGCACCCCCCAACAGAAACATCTTGTGACAGTGAACCTGATGG	133
Oy	593	ATTAGTGGCAGGCTGTGGGGGCCCCCTGGCCCCACACATTAATCGGGGATATAGGCGCACGT	65.2
Db	134	ATTAGTGGCAGGCTGTGGGGGCCCCCTTGCCCCCCACAGATTAATCGGGGATATAGGCGCACGT	133
Oy	653	GCTTACGGCAGCGCGGGGGCGATGAGGAGGGCCGCCACCTAACGCGAGGTATCGGCC	71.2
Db	194	GCTACGGGAGCGGGGGGGCGATGAGGAGGGCGCGCCACCTAACGCGAGGTATCGGCC	253
Oy	713	ACTTACCCGGGGTCTCCCTTCACAGACACGAGACAGTGGGCGGCCCTCTTGTGTGGAGG	77.2
Db	254	ACTTACCCGGGGTCTCTCTTCACAGACACGAGACAGTGGGCGGCCCTCTTGTGTGGAGG	313.3
Oy	773	GGACCCGGCTCCACACACATCATCGGCGCCCTTAGAGAGCGGACCATCTGTGAGGACAAAG	83.2

D <sub>b</sub>	314	GGACCCGCGCTCACACACATCGGCCCTTAGAGCGCACCATTGGAGCAAG	373
O <sub>y</sub>	833	AGAAGGATAAACAAGAAAGACACCCCTGC	861
D <sub>b</sub>	374	AGAAAGATAAACAGAAAGACACCCCTGC	402

```
Search completed: March 18, 2003, 08:14:53
Job time : 1303 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 12:10:39 ; Search time 2671.78 Seconds

(without alignments)  
9378.574 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861  
Sequence: 1 atgcacgcctgtgtgggggt.....aacagaaaggacacccctc 861

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenBml:.\*  
1: gb\_da:.\*  
2: gb\_htg:.\*  
3: gb\_in:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_sts:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vl:.\*  
15: em\_ba:.\*  
16: em\_fun:.\*  
17: em\_hum:.\*  
18: em\_in:.\*  
19: em\_mu:.\*  
20: em\_om:.\*  
21: em\_or:.\*  
22: em\_ov:.\*  
23: em\_pat:.\*  
24: em\_ph:.\*  
25: em\_pl:.\*  
26: em\_ro:.\*  
27: em\_sts:.\*  
28: em\_un:.\*  
29: em\_vl:.\*  
30: em\_htg\_hum:.\*  
31: em\_htg\_inv:.\*  
32: em\_htg\_other:.\*  
33: em\_htg\_mus:.\*  
34: em\_htg\_pln:.\*  
35: em\_htg\_rtd:.\*  
36: em\_htg\_man:.\*  
37: em\_htg\_vrt:.\*  
38: em\_sy:.\*  
39: em\_htgo\_hum:.\*  
40: em\_htgo\_mus:.\*  
41: em\_htgo\_other:.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861	100.0	861	6 AX392419	AX392419 Sequence
2	861	100.0	1321	6 AX392417	AX392417 Sequence
3	861	100.0	4839	6 AF305616	AF305616 Homo sapi
4	839.4	97.5	1061	9 BC015918	BC015918 Homo sapi
5	754.2	87.6	1141	9 AF224278	AF224278 Homo sapi
6	752	87.3	1818	9 AX128643	AX128643 Homo sapi
7	612.4	71.1	878	6 AX392428	AX392428 Sequence
8	546.4	63.5	61505	9 AF305426	AF305426 Homo sapi
9	546.4	63.5	130435	9 HS718J7	AL035541 Human DNA
10	445.6	51.8	693	6 AX392430	AX392430 Sequence
11	421	48.9	651	10 AF220208	AF220208 Mus muscu
12	417.4	48.5	812	6 AX011709	AX011709 Sequence
13	401.4	46.6	408	6 AX071267	AX071267 Sequence
14	358	41.6	156075	2 AL837509	AL837509 Mus muscu
15	352.2	40.9	8494	9 AF009424	AF009424 Homo sapi
16	350	40.7	2170	9 AK055028	AK055028 Homo sapi
17	349	40.5	8093	6 AX392432	AX392432 Sequence
18	349	40.5	8093	6 AF009426	AF009426 Homo sapi
19	341.8	39.7	475	6 AX392431	AX392431 Sequence
20	284.2	33.0	8440	9 AF009425	AF009425 Homo sapi
21	281	32.6	8039	9 AF009427	AF009427 Homo sapi
22	250.2	29.1	172692	2 AP001013	AP001013 Homo sapi
23	250.2	29.1	173709	2 AP001010	AP001010 Homo sapi
24	250.2	29.1	181083	2 AP001268	AP001268 Homo sapi
25	240.8	28.0	766	10 BC022716	BC022716 Mus muscu
26	232.6	27.0	66972	2 AC117364	AC117364 Rattus no
27	231	26.8	183619	2 AC111069	AC111069 Mus muscu
28	228.2	26.5	150542	2 AC097603	AC097603 Rattus no
29	215.8	25.1	167758	2 AC110189	AC110189 Mus muscu
30	157.4	18.3	150224	9 HSJ1059L7	AL121913 Human DNA
31	152.2	17.7	167758	2 AC110189	AC110189 Mus muscu
32	120	13.9	159824	2 AC111878	AC111878 Rattus no
33	62	7.2	125020	9 AF429315	AF429315 Homo sapi
34	61.8	7.2	125020	9 AF429315	AF429315 Homo sapi
35	60.6	7.0	136452	2 AC111699	AC111699 Rattus no
36	59.8	6.9	88203	5 AC097628	AC097628 Rattus no
37	59.4	6.9	152983	2 AC094473	AC094473 Rattus no
38	58.8	6.8	219952	2 AC084804	AC084804 Mus muscu
39	58	6.7	141892	2 AC023197	AC023197 Mus muscu
40	57.8	6.7	26537	2 AC087228	AC087228 Mus muscu
41	56.6	6.6	236562	2 AL772338	AL772338 Mus muscu
42	56.2	6.5	45905	2 AC127121	AC127121 Rattus no
43	56.2	6.5	131364	2 AC128769	AC128769 Rattus no
44	55.8	6.5	300695	2 AC079431	AC079431 Mus muscu
45	55.6	6.5	166900	2 AC130935	AC130935 Rattus no

# ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AX392419	AX392419	Sequence 3 from Patent WO0216416.	AX392419	AX392419.1	GI:19700734	human.	Homo sapiens	1	Lee, R.T., Landschulz, K.T., Kennedy, S.P., Thompson, J.F. and Turi, T.G.	Diagnosis and treatment of cardiovascular conditions
							Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			

us-09-934-249-3.rtf

08/ 663337-9

200 agagagagagatgccctgctcctcagagagatgcttgccctcgagagacacagtctca 210

501 GGCACGGAATCCAGAGCGCGCAGGCTACGCCCGGCTCGGCCACCGAGCCGCTGGCC 360

Db 713 GGCACGGAATCCAGAGCCGCGAGTCTACGCCGCCCTCGGCGCCACCCAGCCCTGCGC 772  
QY GTGCGCGCCTTGCAGCCAGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTG 420  
Db 773 GTGCGCGCCTTGCAGCCAGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTG 832  
QY 421 CAGCAGAGATGAGCTGCGACCCACCATCTGCTGTGAGAGGAGAGAGAGAGAGAGAGAG 480  
Db 833 CAGCAGAGATGAGCTGCGACCCACCATCTGCTGTGAGAGGAGAGAGAGAGAGAGAGAG 892  
QY 481 TACCAAG 540  
Db 893 TACCAAG 952  
QY 541 GAGTGGGAG 600  
Db 953 GAGTGGGAG 1012  
QY 601 AGGCTGGGAG 660  
Db 1013 AGGCTGGGAG 1072  
QY 661 AGGCGGGGGCGATGAG 720  
Db 1073 AGGCGGGGGCGATGAG 1132  
QY 721 GGGTCTCTCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 780  
Db 1133 GGGTCTCTCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1192  
QY 781 CTCCACACACACACACACATGCGCGCCCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 840  
Db 1193 CTCCACACACACACACATGCGCGCCCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1252  
QY 841 AAACAGAAAGAGACCTCTC 861  
Db 1253 AAACAGAAAGAGACCTCTC 1273

RESULT 3  
AF305616 4839 bp mRNA linear PRI 21-OCT-2001

LOCUS AF305616  
DEFINITION Homo sapiens STAG1/PMEP1 mRNA, complete cds.  
ACCESSION AF305616  
VERSION AF305616.1 GI:16303741  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 4839)  
AUTHORS Rae,F.K., Hooper,J.D., Nicol,D.I. and Clements,J.A.  
TITLE Characterization of a novel gene, STAG1/PMEP1, upregulated in renal cell carcinoma and other solid tumors  
Mol. Carcinog. 32 (1), 44-53 (2001)

JOURNAL MEDLINE 21453682  
PUBMED 11568975  
REFERENCE 2 (bases 1 to 4839)  
AUTHORS Rae,F.K., Hooper,J.D., Nicol,D.I. and Clements,J.A.  
TITLE Direct Submission  
JOURNAL Submitted (14-SEP-2000) Centre for Molecular Biotechnology, Queensland University of Technology, 2 George St, Brisbane, QLD 4001, Australia

FEATURES  
Source  
1. 4839  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20q13.2-q13.33"  
321..1184  
/note="unknown function"  
/codon\_start=1  
/product="STAG1/PMEP1"

/protein\_id="AAL16781.1"  
/db\_xref="GI:16303742"  
/translation="MHRMGVNSTFAAAAGAPVNSCTCNCRSLFOSMETTELEPVQI  
IIIVYMMWVVVITCLISYKLSARFISHSQRRERDLSEGLMSESVSN  
GIEPQVYAPRPRLDLAVPPFAQREFRHRQPYPLQHEIDLPITISLSDGEPP  
YQGPCTQLDPEQDLNRSVAPRNRTIFDLDMSARLGPSPSSNSGISATC  
YGSGRMEGPPPTSEVIGHYPGSSFOHOSSGPPSILEGRHLHRTIABLESAIWS  
KERDKRGHPL"

polyA\_signal  
2158..2163  
polyA\_signal  
2463..2468  
polyA\_signal  
4818..4823  
BASE COUNT 1103 a 1313 c 1290 g 1133 t  
ORIGIN

Query Match 100.0%; Score 861; DB 9; Length 4839;  
Best Local Similarity 100.0%; Pred. NO. 4e-136;  
Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGCTTGATGGGGGTCAACAGCACCGCGCGCGCGCGCGAGCCCAATGTC 60  
Db 321 ATGCACCGCTTGATGGGGGTCAACAGCACCGCGCGCGCGCGAGCCCAATGTC 380  
QY 61 TCCTGACAGTGCACACTGCAAAAGCTCTTTGTTCCAGAGCATGAGATCAGGAGCTGGAG 120  
Db 381 TCCTGACAGTGCACACTGCAAAAGCTCTTTGTTCCAGAGCATGAGATCAGGAGCTGGAG 440  
QY 121 TTTGTTCAATCATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
Db 441 TTTGTTCAATCATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 500  
QY 181 CTGCTGAGACCACTACAACTGCTCTGACAGGTCCTTCAATCAGCCGACAGCAGAGAGAGAG 240  
Db 501 CTGCTGAGACCACTACAACTGCTCTGACAGGTCCTTCAATCAGCCGACAGCAGAGAGAGAG 560  
QY 241 AGGAGAGAAGATGCCCTCTCTCAGAGAAGATGCTGTGGCCCTCGGAGAGACACTGTCA 300  
Db 561 AGGAGAGAAGATGCCCTCTCTCAGAGAAGATGCTGTGGCCCTCGGAGAGACACTGTCA 620  
QY 301 GGCACAGGAATCCAGAGACCGCAGGTCTACGCCCGCTCGGCCACCGACCGCTGCGC 360  
Db 621 GGCACAGGAATCCAGAGACCGCAGGTCTACGCCCGCTCGGCCACCGACCGCTGCGC 680  
QY 361 GTGCGCGCCTTTCGCCACAGCGGAGAGCGCTTCCACCGCTTCCAGAGCCAGTATCCGTA 420  
Db 681 GTGCGCGCCTTTCGCCACAGCGGAGAGCGCTTCCACCGCTTCCAGAGCCAGTATCCGTA 740  
QY 421 CAGCAGAGATGAGCTGCGACCCACCATCTGCTGTGAGAGAGGAGAGAGAGAGAGAGAG 480  
Db 741 CAGCAGAGATGAGCTGCGACCCACCATCTGCTGTGAGAGAGGAGAGAGAGAGAGAGAG 800  
QY 481 TACCAAG 540  
Db 801 TACCAAG 860  
QY 541 GAGTGGTGGCGAGCCGCCAAACAGAACATCTTGCAGAGTGAAGTGAAGTGAAGTGAAGT 600  
Db 861 GAGTGGTGGCGAGCCGCCAAACAGAACATCTTGCAGAGTGAAGTGAAGTGAAGTGAAGT 920  
QY 601 AGGCTGGGAG 660  
Db 921 AGGCTGGGAG 980  
QY 661 AGGCGGGGGCGATGAG 720  
Db 981 AGGCGGGGGCGATGAG 1040  
QY 721 GGGTCTCTCTTCCAGACAGCAG 780  
Db 1041 GGGTCTCTCTTCCAGACAGCAG 1100  
QY 781 CTCCACACACACACATGCGCGCCCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 840  
Db 1101 CTCCACACACACACATGCGCGCCCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1160

QY 841 AACAGAAAGACACCCCTC 861  
 Db 1161 AACAGAAAGACACCCCTC 1181  
 RESULT 4  
 BC015918  
 LOCUS  
 DEFINITION Homo sapiens, clone MGC:20374 IMAGE:4559576, mRNA, complete cds.  
 ACCESSION BC015918  
 VERSION BC015918.1 GI:16198474  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 1061)  
 Strausberg, R.  
 Direct Submission  
 Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: DCTP/DRP  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 PC Cancer Agency by: Genome Sequence Centre,  
 InfoDbGsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven Ness, Pawan Pandon, Anna-Lisa Pridhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zynderduyn, Marco Marra.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAL Plate: 29 Row: e Column: 5  
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.  
 Location/Qualifiers  
 1..1061  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:20374 IMAGE:4559576"  
 /tissue\_type="Kidney, renal cell adenocarcinoma"  
 /clone\_lib="NIH MGC\_14"  
 /lab\_host="DH10B-R"  
 /note="vector: pORb7"  
 131..844  
 /codon\_start=1  
 /product="Unknown (protein for MGC:20374)"  
 /protein\_id="AAH15918.1"  
 /db\_xref="GI:16198475"  
 /translation="MMWVWVITCLSHYKISAKSFIRHSGRRREDALSSGCLMP  
 SESTYSGNGIEPPOVYAPRPDLAPPEAFQREHFRPPTPYLQHEIDUPPTSL  
 SDGEPPPYQSGCTQLQNDPEQDLNESPVRAPNRTIFPSDLMSARLGPCPPSS  
 NSGISATCYGSGGMEGPPPTSEVIGHYPPSSQHQSSGPPSLSESTRLHHHIAIP  
 LESAIAWKEKDKMGHPL"  
 BASE COUNT 239 a 350 c 311 g 161 t  
 ORIGIN  
 Query Match 97.5%; Score 839.4; DB 9; Length 1061;  
 Best Local Similarity 99.9%; Pred. No. 2,4e-134;  
 Matches 840; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 CAACAGACCCGCGCGCGCGCGGAGACCCCAATGTCCTGACGTGCAACTGCAA 80  
 Db 1 CAACAGACCCGCGCGCGCGCGCGGAGACCCCAATGTCCTGACGTGCAACTGCAA 60  
 QY 81 AGCGCTTTTGTTCAGAGCATGGAGATACAGGAGCTGGATTGTTCATATCATAT 140  
 Db 61 AGCGCTTTTGTTCAGAGCATGGAGATACAGGAGCTGGATTGTTCATATCATAT 120  
 QY 141 CGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 200  
 Db 121 CGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 QY 201 GTCTGACAGGTCTTATATACACCGGACACAGGAGGAGAGAGAGATGCTGTCT 260  
 Db 181 GTCTGACAGGTCTTATATACACCGGACACAGGAGGAGAGAGAGATGCTGTCT 240  
 QY 261 CTCAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320  
 Db 241 CTCAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 321 GCAGGTCTACGCGCGCGCTGCGCCACGACGCGCTGCGCGCTGCGCGCTGCGCGCG 380  
 Db 301 GCAGGTCTACGCGCGCGCTGCGCCACGACGCGCTGCGCGCTGCGCGCTGCGCGCG 360  
 QY 381 GGAGCGCTTCACCGCTTCACCGCTTCACCGCTTCACCGCTTCACCGCTTCACCGCT 440  
 Db 361 GGAGCGCTTCACCGCTTCACCGCTTCACCGCTTCACCGCTTCACCGCTTCACCGCT 420  
 QY 441 ACCCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500  
 Db 421 GCCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 501 CCAGCTTGGGAGACCCGAGACAGCTGGAATGAACGGAGTGGTGGCGGACCCCG 560  
 Db 481 CCAGCTTGGGAGACCCGAGACAGCTGGAATGAACGGAGTGGTGGCGGACCCCG 540  
 QY 561 AAACAGAACATCTTGCAGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 620  
 Db 541 AAACAGAACATCTTGCAGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 600  
 QY 621 CCCAGCACTAAGTGGGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 680  
 Db 601 CCCAGCACTAAGTGGGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
 QY 681 GCCGCGCGCCACTTACAGGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 740  
 Db 661 GCCGCGCGCCACTTACAGGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
 QY 741 GCAGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 800  
 Db 721 GCAGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
 QY 801 GCCCTTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 860  
 Db 781 GCCCTTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840  
 QY 861 C 861  
 Db 841 C 841  
 RESULT 5  
 AF224278 1141 bp mRNA linear PRI 18-JUL-2000  
 LOCUS  
 DEFINITION Homo sapiens PMP21 protein (PMP21) mRNA, complete cds.  
 ACCESSION AF224278  
 VERSION AF224278.1 GI:9255808  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.



	REFERENCE	1 (bases 1 to 1141)
AUTHORS	Xu,L.L., Shammugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G., Moul,J.W. and Srivastava,S.	
TITLE	A novel androgen-regulated gene, PMEPA1, located on chromosome 20q13 exhibits high level expression in prostate	
JOURNAL	Genomics 66 (3), 257-263 (2000)	
MEDLINE	20334621	
PUBMED	10873380	
REFERENCE	2 (bases 1 to 1141)	
AUTHORS	Xu,L.L., Shammugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G., Moul,J.W. and Srivastava,S.	
JOURNAL	Direct Submission	
FEATURES	Submitted (12-JAN-2000) CDR, USUHS, 1530 East Jefferson Street, Rockville, MD 20852, USA	
SOURCE	Location/Qualifiers 1..1141 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="20" /map="20q13.31-13.33" /cell_line="LNCap" 1..1141 /gene="PMEPA1" 96..854 /gene="PMEPA1" /note="Type Ib transmembrane protein; expression is induced in response to the synthetic androgen, R1881; expression is abundant in, and restricted to, prostatic glandular epithelial cells; similar to the predicted protein encoded by sequence deposited at GenBank Accession Number NP_004329" /codon_start=1 /product="PMEPA1 protein" /protein_id="AAF66322.1" /db_xref="GI:9255809" /translation="MADEEFYQIIIVVWVVVTCLLSHKLSARSPISHSOGRRRRDAUSSECTLMPSVESTVSNGIPEPVYAPRPDRLAVPFADERHRPOPTTYRYLDARIGDLPPITSLSDGEPPYPQGCTQLRLPDQLELNRESVAPRNRTLFDSBLMDARLEGPCPPSSNSGISATCYGSGGSRMGPPPTSEVIGHYPASSFQHQQSGPPSLTEGTNRHHHHPLESAAIMSKENOKRHP.L"	
CDS		
gene		
BASE COUNT	271 a 350 c 336 g 184 t	
ORIGIN		
Query Match	87.6%; Score 754.2; DB 9; Length 1141;	
Best Local Similarity	98.3%; Pred. No. 9.3e-120;	
Matches 762; Conservative	0; Mismatches 13; Indels 0; Gaps 0;	
OY	87 TTTGTTCCAGACATGGAGATCAGGAGCTGGAGATTGTTGTCAGATCATCATCTGCTGT 146	
DB	77 TCTCCTGGCAACCAGGCACATGCGGAGCTGGAGTGTGTTGTCAGATCATCATCTGCTGT 136	
OY	147 GGTCATGATGGTGATGGTGGTGGTGAATCACCGGCCCTGGAGACCACTAACGCTGTGC 206	
DB	137 GGTGATGATGGTGAATGGTGGTGGTGAATCACGCTGCTGTAACCACTGTCTGC 196	
OY	207 ACGGTCCTTCATCAGCCGCGACAGCCAGGCGGCGAGAGAAGAAGATGCCCTGCTCAGA 266	
DB	197 ACGGTCCTTCATCAGCCGCGACAGCCAGGCGGCGAGAGAAGAAGATGCCCTGCTCAGA 256	
OY	267 AGGATGCTGTGGCCCTGGGAGAGCACAAGTGTCCAGGCAACGGAATCCAGAGCCGAGGT 326	
DB	257 AGGATGCTGTGGCCCTGGGAGAGCACAAGTGTCCAGGCAACGGAATCCAGAGCCGAGGT 316	
OY	327 CTACGCCCCGCTCGGCGCCACGACCGCTGGCGTGGCGGCCCTTCGGCCACGGGAGAG 386	
DB	317 CTACGCCCCGCTCGGCGCCACGACCGCTGGCGTGGCGGCCCTTCGGCCACGGGAGAG 376	
OY	387 CTTCACGCGCTTCACAGCCACCTATTCGATTCCTGAGACAGAGATGACTGGCCACCCAC 446	
DB	377 CTTCACGCGCTTCACAGCCACCTATTCGATTCCTGAGACAGAGATGACTGGCCACCCAC 436	
OY	447 CATCTCGCTGTAGACGGGAGAGAGCCCAACCTTACCAAGGGCCCCTGCACTCTCAGCT 506	
DB	437 CATCTCGCTGTAGACGGGAGAGAGCCCAACCTTACCAAGGGCCCCTGCACTCTCAGCT 506	

[illegible]



```

LOCUS       AF305426                61505 bp    DNA    linear    PRI 12-OCT-2001
DEFINITION   Homo sapiens solid tumor-associated 1 protein (STAG1/PMEPA1) gene,
ACCESSION   AF305426
VERSION     AF305426.1    GI:15824468
KEYWORDS
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 61505)
AUTHORS     Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.
TITLE       Identification and characterization of a novel gene, STAG1,
            up-regulated in renal cell carcinoma and other solid tumours
            Unpublished
JOURNAL     2 (bases 1 to 61505)
REFERENCE   Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.
AUTHORS
TITLE       Direct Submission
JOURNAL     Submitted (13-SEP-2000) Centre for Molecular Biotechnology,
            Queensland University of Technology, 2 George St, Brisbane, QLD
            4001, Australia

FEATURES
source      location/Qualifiers
            1..61505
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="20q13.2-q13.33"
            1..61505
            /gene="STAG1/PMEPA1"
            join(1..429,50206..50360,56817..56870,57305..61505)
            /gene="STAG1/PMEPA1"
            /product="solid tumor-associated 1 protein"
            join(321..429,50206..50360,56817..56870,57305..57850)
            /gene="STAG1/PMEPA1"
            /codon_start=1
            /product="solid tumor-associated 1 protein"
            /protein_id="AL09357.1"
            /db_xref="GI:15824469"
            /translation="MHRIMGVNSTAAAGOPVSGTCNCKRSLRSMETLEFVOI
            IITVYMMVWVYITCLSHYKLSARSPISRHSGRRDALSSSECLMPSESTVGN
            GIDPEQVIAVPRFDRLAVPEFAQRERFRFQPTPYLQHEIDLPTTISLSDSEPP
            YOGPCYLOLRDPEQOLELNRESVRAPNRTIFDSDLSARLGGPCPPSSNSISATC
            YGSGGMEGPPPTYSVIVGHYPPSSSFQHOQSSGPPSLLESTRLHHTHIALPESAAIMS
            KERDKOKGHPVL"

BASE COUNT   13329 a 16074 c 17430 g 14672 t
ORIGIN
Query Match      63.5%; Score 546.4; DB 9; Length 61505;
Best Local Similarity 98.9%; Pred. No. 2,4e-84;
Matches 550; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 306 CGGAATCCAGAGCCGAGTCTACGCCGCCCTCGGCCACGACGCGCTGGCCGCTGCC 365
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 57292 CTGCTTCCTCCAGCGGAGGTCTACGCCGCCCTCGGCCACGACGCGCTGGCCGCTGCC 57351

QY 366 GCCCTTGGCCAGGAGGAGGCTTCACACGCTTCACGCCACCACTATCGCTGACGACA 425
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 57352 GCCCTTGGCCAGGAGGAGGCTTCACACGCTTCACGCCACCACTATCGCTGACGACA 57411

QY 426 CGAGATGACCTGCACCCACATCTGCTGTACAGAGGGGAGAGAGGCCACCCCTACCA 485
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 57412 CGAGATGACCTGCACCCACATCTGCTGTACAGAGGGGAGAGAGGCCACCCCTACCA 57471

QY 486 GGGGCCCTGCACCTTCAGCTTGGGAGACCCGAGACAGCAGCTGGAAGTGAACGGGAGTC 545
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 57472 GGGGCCCTGCACCTTCAGCTTGGGAGACCCGAGACAGCAGCTGGAAGTGAACGGGAGTC 57531

QY 546 GGTCGGCGACCCCAAGAACATCTTGAGAGTGAAGTGAATGATGTCAGAGCT 605
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 57532 GGTCGGCGACCCCAAGAACATCTTGAGAGTGAAGTGAATGATGTCAGAGCT 57591

QY 606 GGGCGGCCCTGCCCCCGACAGCACTGAAGCGGAGTGAAGCGGAGCGG 665
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

Db 57592 GGGCGGCCCTGCCCCCGACAGCACTGAAGCGGAGTGAAGCGGAGTGAAGCGGAGCGG 57651
QY 666 CGGGCGCATGAGAGGGGGCGCCGCTACACTACAGGAGGTATGAGGCACACTACCGGGGCTC 725
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 57652 CGGGCGCATGAGAGGGGGCGCCGCTACACTACAGGAGGTATGAGGCACACTACCGGGGCTC 57711

QY 726 CTGCTTCACAGCAGAGAGAGAGTGGGCGCCCTCTGCTGAGAGGGAGCCCGGCTCCA 785
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 57712 CTGCTTCACAGCAGAGAGAGTGGGCGCCCTCTGCTGAGAGGGAGCCCGGCTCCA 57771

QY 786 CCACACACATGCGCGCCCTAGAGAGCGAGCATCTGAGCAAAAGAGAGATTAACA 845
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 57772 CCACACACATGCGCGCCCTAGAGAGCGAGCATCTGAGCAAAAGAGAGATTAACA 57831

QY 846 GAAAGGACACCCCTCTC 861
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 57832 GAAAGGACACCCCTCTC 57847

RESULT 9
LOCUS       HS718U7/c
DEFINITION   Human DNA sequence from clone RP4-718U7 on chromosome
            20q13.31-13.33 Contains the PKI1 gene for soluble
            phosphoenolpyruvate carboxylkinase 1, part of a novel gene similar
            to mouse DLM-1 (tumour stroma and activated macrophage protein),
            the 3' end of the TMEMPA1 gene encoding an androgen induced 1b
            transmembrane protein (PMEPA1), two putative novel genes, a CPG
            island, ESTs, STS and GSSs, complete sequence.
            AL035541
            AL035541.15 GI:11546043
            HTG: CPG island: DLM-1; macrophage protein: PKI1;
            phosphoenolpyruvate carboxylkinase; PMEPA1; TMEMPA1; transmembrane
            protein.
            human.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 130435)
REFERENCE   Direct Submission
AUTHORS     Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
            requests: clonerequests@sanger.ac.uk
            On Dec 5, 2000 this sequence version replaced gi:10198628.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            This sequence is the entire insert of clone RP4-718U7 The true left
            end of clone RP5-1007E6 is at 71437 in this sequence. The true
            right end of clone RP4-579F20 is at 43945 in this sequence.
            RP4-718U7 is from the library RCT-4 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the

```

```
FEATURES      assembly was confirmed by restriction digest.
source         Location/Qualifiers
1. 130435
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="20"
   /map="q13.31-q13.33"
   /clone="RP4-718J7"
   /clone_11b="RPC1-4"
   319..478
repeat_region /note="5 copies 32 mer 67% conserved"
               370..463
repeat_region /note="2 copies 47 mer 87% conserved"
               427..626
repeat_region /note="4 copies 50 mer 71% conserved"
               707..812
repeat_region /note="53 copies 2 mer cc 61% conserved"
               967..2752
repeat_region /note="893 copies 2 mer 99 54% conserved"
               982..2730
repeat_region /note="33 copies 53 mer 54% conserved"
               1177..1820
repeat_region /note="4 copies 161 mer 64% conserved"
               1204..1371
repeat_region /note="3 copies 56 mer 75% conserved"
               1358..2757
repeat_region /note="28 copies 50 mer 54% conserved"
               1675..2718
repeat_region /note="18 copies 58 mer 55% conserved"
               1843..2346
repeat_region /note="9 copies 56 mer 64% conserved"
               1852..2079
repeat_region /note="4 copies 57 mer 86% conserved"
               2028..2632
repeat_region /note="11 copies 55 mer 60% conserved"
               2087..2257
repeat_region /note="3 copies 57 mer 83% conserved"
               2228..2497
repeat_region /note="5 copies 54 mer 75% conserved"
               2583..2716
repeat_region /note="2 copies 67 mer 82% conserved"
               3378..3426
repeat_region /note="LIMB4 repeat: matches 6088..6136 of consensus"
               3799..3896
repeat_region /note="Charliel repeat: matches 681..781 of consensus"
               5331..5793
repeat_region /note="MUR1B repeat: matches 14..466 of consensus"
               5797..5988
repeat_region /note="6 copies 32 mer 86% conserved"
               5855..5978
repeat_region /note="31 copies 4 mer gcac 61% conserved"
               6474..6591
repeat_region /note="MIR repeat: matches 91..218 of consensus"
               6592..6723
repeat_region /note="FLAM_C repeat: matches 1..132 of consensus"
               6952..7021
repeat_region /note="L2 repeat: matches 2637..2705 of consensus"
               7358..7671
repeat_region /note="AlusX repeat: matches 1..312 of consensus"
               8521..8554
repeat_region /note="Alu repeat: matches 1..34 of consensus"
               8863..9217
repeat_region /note="MUR1D repeat: matches 105..505 of consensus"
               9228..9746
repeat_region /note="L2 repeat: matches 1459..2028 of consensus"
               9880..10089
repeat_region /note="L2 repeat: matches 2292..2511 of consensus"
               10094..10206
repeat_region /note="AluYb repeat: matches 188..300 of consensus"
               10263..10573
repeat_region /note="AluYb repeat: matches 1..308 of consensus"
               10658..10753
repeat_region /note="L2 repeat: matches 2647..2749 of consensus"
               11816..12380
repeat_region /note="LTR19B repeat: matches 1..580 of consensus"
               12519..12813
repeat_region /note="LIMD2 repeat: matches 6032..6331 of consensus"
               13539..13726
repeat_region /note="3 copies 56 mer 73% conserved"
               13585..13716
repeat_region /note="22 copies 6 mer ctctct 67% conserved"
               13586..13717
repeat_region /note="66 copies 2 mer tc 68% conserved"
               13588..13715
repeat_region /note="4 copies 32 mer 78% conserved"
               13589..13724
repeat_region /note="34 copies 4 mer cttt 77% conserved"
               complement(13622..14142)
misc_feature /note="match: GSS: Em:AO592603"
               14313..14427
repeat_region /note="LTR16C repeat: matches 257..386 of consensus"
               15015..15110
repeat_region /note="3 copies 32 mer 79% conserved"
               15021..15110
repeat_region /note="45 copies 2 mer ac 80% conserved"
               15026..15109
repeat_region /note="7 copies 12 mer 82% conserved"
               15028..15111
repeat_region /note="14 copies 6 mer cacaca 82% conserved"
               15029..15108
repeat_region /note="20 copies 4 mer acac 83% conserved"
               15273..15399
repeat_region /note="L2 repeat: matches 2159..2285 of consensus"
               15420..15721
repeat_region /note="AlusX repeat: matches 1..303 of consensus"
               16333..16396
repeat_region /note="2 copies 32 mer 98% conserved"
               16575..16690
repeat_region /note="L2 repeat: matches 2596..2711 of consensus"
               18049..18169
repeat_region /note="MIR repeat: matches 86..211 of consensus"
               18312..18438
repeat_region /note="MIR repeat: matches 138..250 of consensus"
               18585..18776
repeat_region /note="MER20 repeat: matches 7..217 of consensus"
               19211..19290
repeat_region /note="MIR repeat: matches 72..154 of consensus"
               20413..20463
repeat_region /note="MIR repeat: matches 90..143 of consensus"
               complement(21005..21672)
misc_feature /note="match: GSS: Em:AO748384"
               complement(21187..21672)
misc_feature /note="match: GSS: Em:AO776209"
               complement(21278..21663)
misc_feature /note="match: GSS: Em:AO136459"
               complement(21392..21672)
misc_feature /note="match: GSS: Em:AO892114"
               21667..21775
repeat_region /note="MIR repeat: matches 73..192 of consensus"
               23577..23646
repeat_region /note="2 copies 35 mer 100% conserved"
               24257..24591
repeat_region /note="L2 repeat: matches 2336..2710 of consensus"
               25580..25611
repeat_region /note="16 copies 2 mer tc 87% conserved"
               26333..26643
repeat_region /note="AluY repeat: matches 1..310 of consensus"
               27603..27684
repeat_region /note="L2 repeat: matches 2661..2739 of consensus"
               27823..28041
repeat_region /note="MIR repeat: matches 38..242 of consensus"
               28129..28259
repeat_region /note="FLAM_C repeat: matches 1..127 of consensus"
               31203..31356
repeat_region /note="LIMB5 repeat: matches 6015..6176 of consensus"
               32224..32522
```

LOCUS	AX392430	693 bp	DNA	linear	PAT 23-MAR-2002
DEFINITION	Sequence 14 from Patent WO0216416.				
ACCESSION	AX392430				
VERSION	AX392430.1	GI:19700746			
KEYWORDS	.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and Turi,T.G.				
TITLE	Diagnosis and treatment of cardiovascular conditions				
JOURNAL	Patent: WO 0216416-A 14 28-FEB-2002;				
FEATURES	THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)				
Source	Location/Qualifiers				
	1..693				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	101 a	205 c	237 g	149 t	1 others
ORIGIN					
Query Match	51.8%, Score 445.6; DB 6; Length 693;				

```

AF220208 RESULT 11
AF220208 LOCUS 651 bp mRNA linear ROD 05-MAR-2001
DEFINITION Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.
ACCESSION AF220208
VERSION AF220208.1 GI:12004973
KEYWORDS
SOURCE
ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 651)
Jolliffe,C.N., Harvey,K.F., Haines,B.P., Parasivam,G. and Kumar,S.
Identification of multiple proteins expressed in murine embryos as
binding partners for the WW domains of the ubiquitin-protein ligase
Nedd4
Biochem. J. 351 Pt 3, 557-565 (2000)
JOURNAL
MEDLINE
20498735
PUBMED
11042109
REFERENCE
2 (bases 1 to 651)
Jolliffe,C.N. and Kumar,S.
Direct Submission
Submitted (30-DEC-1999) Division of Haematology, Hanson Centre for
Cancer Research, IMVS, Frome Rd, Adelaide, SA 5000, Australia
JOURNAL
TITLE
Location/Qualifiers
1. 651
FEATURES
SOURCE
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="Novagen Cat. No. 69640-3"
/dev_stage="16 day embryo"
<1. .603
/note="N4MBP4"
/codon_start=1

```

Query Match	48.9%	Score 421;	DB 10;	Length 651;
Best Local Similarity	83.1%;	Pred. No. 1.2e-62;		
Matches 526; Conservative	0;	Mismatches 90;		

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

RESULT	12
AX011709	
LOCUS	
DEFINITION	AX011709 812 bp DNA
ACCESSION	Sequence 107 from Patent WO955858.
VERSION	AX011709
KEYWORDS	AX011709.1 GI:9998233
SOURCE	
ORGANISM	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 812)
AUTHORS	Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and Pilatky,C.

TITLE	JOURNAL	FEATURES	SOURCE	BASE COUNT	ORIGIN
Human nucleic acid sequences obtained from pancreas tumor tissue	Patent: WO 955858-A 107 04 NOV-1999.	SCMITT ARMIN (DE); SPECHT THOMAS (DE); BERND (DE); ROSENTHAL ANDRE (DE); PILANSKI CHRISTIAN (DE)	Location/Qualifiers 1..812	157 a 272 c 237 g 146 t	
			/organism="Homo sapiens" /db_xref="taxon:9606"		

Query March	48.5%;	Score 417.4;	DB 6;	Length 812;
Best Local Similarity	92.0%;	Pred. No. 5e-62;		
Matches 461;	Conservative	0;	Mismatches 16;	Indels 24;
				Gaps 1

QY 47 GGCAGCCCAATGTCCTCCAGCTGCACTGCAACGCTCTTTGTTCCAGACATGGAGA 106  
 Db 324 GCCAGCCCAATGTCCTCCAGCTGCACTGCAACGCTCTTTGTTCCAGACATGGAGA 383  
 QY 107 TCACGAGACTGGAGATTGTTTCAGATCATCATGCTGGTGGTATGATGATGGTGG 166  
 Db 384 TCACGAGAGCTGGAGATTGTTTCAGATCATCATGCTGGTGGTGGTGGTGGTGGTGG 425  
 QY 167 TGGTGAATCAGACTCCCTGTGAGCCACTCAAGACTGTCTGCACGGTCCCTTCATGAGCCGC 226  
 Db 426 -----TCAGTCCCTGCTGAGCCACTCAAGACTGTCTGCACGGTCCCTTCATGAGCCGC 479  
 QY 227 ACAGCCAGGGGCGGAGAGAGAAAGATGCCCTGCTCTGAGAAGATGGCTGTGGCCCTCGG 286  
 Db 480 ACAGCCAGGGGCGGAGAGAGAAAGATGCCCTGCTCTGAGAAGATGGCTGTGGCCCTCGG 539  
 QY 287 AGACACAGACTGTTCAGGCAACGGGAATCCCAAGAGCCGCGAGCTCTACGCCCCGCTTCGAGCCA 346  
 Db 540 AGACACAGACTGTTCAGGCAACGGGAATCCCAAGAGCCGCGAGCTCTACGCCCCGCTTCGAGCCA 599  
 QY 347 CGAGCCGCTGCGCTGCGGCCCTCTTGCCGACAGCGGAGCCCTTCCAGCGCTTCCAGGCCA 406  
 Db 600 CGAGCCGCTGCGCTGCGGCCCTTTCGCCGACGGGAGAGCCCTTCCAGCGCTTCCAGGCCA 659  
 QY 407 CCTATCCGACTGTGAGACAGAGATTCACCTGCGACCCACACATCTTCGCTGTGACAGGGGG 466  
 Db 660 CCTATCCGACTGTGAGACAGAGATTCACCTGCGGCCACACATCTTCGCTGTGACAGGGGG 719  
 QY 467 AGGAGCCCCCAACCTTACACAGGGGCCCTCTGCACCTTCAGATTGGGGAGCCCCAGACGACGAC 526  
 Db 720 AGGAGCCCCCAACCTTACACAGGGGCCCTCTGCACCTTCAAGATTGGGGAGCCCCAGGAGGAGT 779  
 QY 527 TGGACATCAACCGGGAGTCGG 547  
 Db 780 TGGAAATTGAACGGGGATGG 800

RESULT 13	AX071267	408 bp	DNA	linear	PAT 25-JAN-2001
LOCUS	AX071267				
DEFINITION	Sequence 1739 from Patent WO0102568.				
ACCESSION	AX071267				
VERSION	AX071267.1	GI:12581618			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 408)				
	Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J., Kassam,A., Reinhard,C., Rendazzo,F., Kennedy,G.C., Pot,D., Lamson,G., Drmanac,R., Girenkov,R., Drmanac,S., Dickson,M., Labat,I., Leshkovitz,D., Kita,D., Garcia,V. and Strache-Crain,B.				
TITLE	Human genes and gene expression products				
JOURNAL	Patent: WO 0102568-A 1739 11-JAN-2001;				
FEATURES	CHIRON CORPORATION (us); HYSEQ, INC. (us)				
	Location/Qualifiers				



```

Db 134299 TTCCAGACAGCAGAAATACCGGCGCATCTCCCTGCTAGAGGGGACCGGCTCCATCAC 134240
QY 790 ACACACATCGGGCCCTAGAGAGCGGCGCATCTGAGAGAGAGATMAAACAGAAA 849
Db 134239 TCGCAGATTTGGCCCACTGGA-----GACAAAGAGAGAGAGAGAGAGAGAAA 134195
QY 850 GGACACCCCTCTC 861
Db 134194 GGTACCCCTCTC 134183

RESULT 15
AF009424 8494 bp mRNA linear PRI 17-FEB-1998
LOCUS AF009424
DEFINITION Homo sapiens clone 22 mRNA, alternative splice variant alpha-1,
ACCESSION AF009424
VERSION AF009424.1 GI:2271468
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 8494)
AUTHORS Yoshikawa,T., Sanders,A.R., Esterling,L.E., Overhauser,J.,
Garner,J.A., Lennon,G., Grewal,R. and Detera-Wadleigh,S.D.
TITLE Isolation of chromosome 18-specific brain transcripts as positional
candidates for bipolar disorder
Am. J. Med. Genet. 74 (2), 140-149 (1997)
JOURNAL MEDLINE 97275951
PUBMED 9129712
REFERENCE 2 (bases 1 to 8494)
AUTHORS Yoshikawa,T., Sanders,A.R., Esterling,L.E. and Detera-Wadleigh,S.D.
TITLE Multiple transcriptional variants and RNA editing in C18orf1, a
novel gene with IDRA and transmembrane domains on 18p11.2
Genomics 47 (2), 246-257 (1998)
JOURNAL MEDLINE 98140124
PUBMED 9479497
REFERENCE 3 (bases 1 to 8494)
AUTHORS Yoshikawa,T. and Detera-Wadleigh,S.D.
TITLE Direct Submission
Submitted (20-JUN-1997) Clinical Neurogenetics Branch, National
Institute of Mental Health, Bethesda, MD 20892, USA
JOURNAL
FEATURES
source Location/Qualifiers
1..8494
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18p11.2"
470..1390
/note="alternatively spliced; alpha-1 form; possible
membrane-spanning protein"
/product="clone 22"
/codon_start=1
/db_xref="GI:2271469"
/db_xref="AF009424"
/translation="MPEAGFOATNAFTGCKTCTSGKCLYLGLVLCNQNDGNSDE
ENCLVLTPEPPGLFNSLEFAQIIIVVTVVWVYVCLLNHYKSTRSFINRPNQ
SRREDGLPQEGCLPNSDAAPRLGASEIMAPSRDFTAPSFQIORDRFSRPOPTP
YVOHEIDLPPTSLSDGEPPPYOGPTCLDRPEQOMELNRESVRAAPNRTIFDSOL
IDIAVYSGPCPPSPNSGSIASGSCNRMGEPPTTSEVGHHPGASFLHQRSMNH
RGRLOPQONNAESTIVPIKGRKRPGLV"

BASE COUNT 2285 a 1900 c 1907 g 2400 t 2 others
ORIGIN
Query Match 40.9%; Score 352.2; DB 9; Length 8494;
Best Local Similarity 73.1%; Pred. No. 5.7e-51;
Matches 482; Conservative 0; Mismatches 166; Indels 9; Gaps 2;
QY 94 CAGACATGAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCATGCTGTGATG 153
Db 635 CCGGCACTCTCACTCGAGAGCTGGAGTTCCGCCAATCATCATCATGCTGTGATG 694

```

```

QY 154 ATGCTATGCTGCTGATCACTGCTCTGAGCCACTACAGCTGTCTGACAGCTCC 213
Db 695 ACGTATGCTGCTGCTGATCACTGCTCTGAGCCACTACAGCTGTCTGACAGCTCC 754
QY 214 TTTCATCAGCGGCGGAGGCGGAGGAGAGAGATGCGCTGCTTCAGAGAGATGC 273
Db 755 TTTCATCAGCGGCGGAGGCGGAGGAGAGAGATGCGCTGCTTCAGAGAGATGC 814
QY 274 CTGTGCGCTTCGAGAGCAGTGTGAGGCAACGGAATCCAGAGCGGAGTTCAGGC 333
Db 815 CTGTGCGCTTCGAGAGCAGTGTGAGGCAACGGAATCCAGAGCGGAGTTCAGGC 868
QY 334 CGGCTTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 393
Db 869 GCGGCTTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 928
QY 394 CGGCTTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 453
Db 929 CGGCTTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 988
QY 454 CTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 513
Db 989 CTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1048
QY 514 CCGGAGCAGCAGCTGGAAGTGAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 573
Db 1049 CCGGAGCAGCAGCTGGAAGTGAACCGGAGGAGGAGGAGGAGGAGGAGGAGG 1108
QY 574 TTTCAGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
Db 1109 TTTCAGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1168
QY 631 AACTGGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 690
Db 1169 AACTGGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1228
QY 691 AACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 749
Db 1229 AACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1287

```

Search completed: March 17, 2003, 15:53:56  
Job time : 2748.78 secs





PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.  
 PA (PRTZ) PFIZER INC.

Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;  
 DR WPI: 2002-280912/32.

DR P-PSDB: AAU78231.

XX Novel nucleic acid molecule encoding Mechanically Induced Vascular  
 XX Receptor-1 polypeptide, useful for treating cardiovascular diseases  
 PS Claim 2: Page 87-88; 105pp; English.

CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having  
 CC cardiac cell anti-apoptotic activity and fragments of it provided  
 CC they are not identical to Genbank sequences A1761441.1, A1594390,  
 CC NM\_004338 and A0177461. Also included are expression vectors, host  
 CC cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of  
 CC MIVR-1, contacting a molecule having cardiac cell anti-apoptotic activity  
 CC with a candidate agent, where the molecule is a nucleic acid molecule  
 CC comprising MIVR-1, IEX-1, VDU-1, Brg-2 and R15-11d or its  
 CC expression product, determining if the anti-apoptotic activity is  
 CC modulated and thereby identifying a modulator. The cardiac cell anti-  
 CC apoptotic molecules and nucleic acids of the invention are useful for  
 CC treating, diagnosing and monitoring progression of such diseases and  
 CC disorders as characterized by increased apoptotic cell-death of vascular  
 CC endothelial cells e.g. cardiac hypertrophy, myocardial infarction,  
 CC stroke, arteriosclerosis and heart failure. The present sequence  
 CC encodes human MIVR-1.

Sequence 1321 BP; 223 A; 493 C; 440 G; 165 T; 0 other:

Query Match 100.0%; Score 861; DB 24; Length 1321;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
 Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCGCTGATGAGGGGTCACACACCGCCGCCGCCGGGAGCCCAATGTC 60  
 DB 413 ATGACCGCTGATGAGGGGTCACACACCGCCGCCGCCGGGAGCCCAATGTC 60  
 QY 61 TCCTGACAGTGAACGAGCGCTTGTTCAGACAGTGAAGATGAGAGCTGAG 120  
 DB 473 TCCTGACAGTGAACGAGCGCTTGTTCAGACAGTGAAGATGAGAGCTGAG 120  
 QY 121 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 532  
 DB 533 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 532  
 QY 181 CTGCTGAGGCACTACAGCTGCTGACAGGTCCTTCATCAGCCGACAGCCAGGGCGG 240  
 DB 593 CTGCTGAGGCACTACAGCTGCTGACAGGTCCTTCATCAGCCGACAGCCAGGGCGG 240  
 QY 241 AGGAGAGAGAGTCCCTCTCTCAGAAAGATGCTGTGGCCCTCGAGAGACAGTGA 300  
 DB 653 AGGAGAGAGAGTCCCTCTCTCAGAAAGATGCTGTGGCCCTCGAGAGACAGTGA 300  
 QY 301 GGCAAGAGAGTCCAGAGCCGAGGCTAGGCCCGGCTTGCCCAACGAGCCCTGAGCC 360  
 DB 713 GGCAAGAGAGTCCAGAGCCGAGGCTAGGCCCGGCTTGCCCAACGAGCCCTGAGCC 360  
 QY 361 GTGCGGCGCTTGGCCGAGGAGCGGCTTCACAGCCGCTTCAGAGCCGCTTCAGT 420  
 DB 773 GTGCGGCGCTTGGCCGAGGAGCGGCTTCACAGCCGCTTCAGAGCCGCTTCAGT 420  
 QY 421 CAGCAGAGATGACCTGACACCAACATCTGCTGTGACAGGGGAGAGCCCAACCC 480  
 DB 833 CAGCAGAGATGACCTGACACCAACATCTGCTGTGACAGGGGAGAGCCCAACCC 480  
 QY 481 TACCAAGGCGCCCTGACACCTTCCAGCTTGGAGAGCCCGAGCAGCAGTGAACCGG 540  
 DB 893 TACCAAGGCGCCCTGACACCTTCCAGCTTGGAGAGCCCGAGCAGCAGTGAACCGG 540

QY 541 GAGTCGAGGCGCGACCCCAACAGACATCTTGACAGTGAAGCTGATGATGTC 600  
 DB 953 GAGTCGAGGCGCGACCCCAACAGACATCTTGACAGTGAAGCTGATGATGTC 600  
 QY 601 AGGCTGGGCGCGCCCGCCCGCCCGCCAGCACTAATCTGGGATCAGCCAGCTGACGGC 660  
 DB 1013 AGGCTGGGCGCGCCCGCCCGCCCGCCAGCACTAATCTGGGATCAGCCAGCTGACGGC 660  
 QY 661 AGGCGGCGCGATGAGAGGGCGCGCCCACTACAGCGAGTCACTGCGCACTACCG 720  
 DB 1073 AGGCGGCGCGATGAGAGGGCGCGCCCACTACAGCGAGTCACTGCGCACTACCG 720  
 QY 721 GGGTCCTCTTCCACACAGCAGAGCACTGCGCGCCCTCTCTGAGAGGGAGCCCG 780  
 DB 1133 GGGTCCTCTTCCACACAGCAGAGCACTGCGCGCCCTCTCTGAGAGGGAGCCCG 780  
 QY 781 CTCACACACACACATGCGCCCTAGAGAGCGCAGCAGCTGAGAGAGAGAGAT 840  
 DB 1193 CTCACACACACACATGCGCCCTAGAGAGCGCAGCAGCTGAGAGAGAGAGAT 840  
 QY 841 AACAGAAAGAGACACCTCTC 861  
 DB 1253 AACAGAAAGAGACACCTCTC 1273

# RESULT 2

AAAT5151  
 ID AAAT5151 standard; cDNA; 969 BP.

AC AAAT5151;

DT 15-JAN-2001 (first entry)

DE cDNA encoding a human TANGO 261 polypeptide.

XX TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hemolytic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW cardiovascular disorder; spleen associated disease; renal disorder;  
 KW brain herniation; latrogenic disease; inflammation; meningitis;  
 KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 6..764

FT /tag- a

FT /product- "TANGO 261"

FT sig-peptide 6..89

FT mat-peptide 90..764

FT /tag- b

FT /tag- c

PN W0200052022-A1.

XX 08-SEP-2000.

PD 01-MAR-2000; 2000WO-US05226.

PF 01-MAR-1999; 990S-0122458.

XX (MILL-) MILLENNIUM PHARM INC.

PA Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

PI WPI: 2000-579269/54.

DR P-PSDB: AAB18449.

XX Novel human and murine secreted proteins designated TANGO 216, 261,

PT 262, 266 and 267 useful as modulating agents of cellular processes,  
XX e.g. for treating cancer -  
PS Claim 2; Fig 5; 175pp; English.

claim 2; fig 3; 1/cpp; English.  
XX

CC The present sequence encodes a human TANGO 261 polypeptide. The  
CC specification also describes TANGO 266, TANGO 216, TANGO 262, and  
CC TANGO 267. The TANGO polypeptides can be used to modulate cellular  
CC proliferation, modulate cellular differentiation and/or modulate  
CC cellular adhesion. The proteins can be used to treat any von Willebrand  
CC factor-associated disorder, regulate extracellular matrix structuring,  
CC cellular adhesion, and cell trafficking and/or migration, modulate  
CC cellular interactions, modulate cell adhesion in proliferative  
CC disorders, such as cancer, modulate the proliferation, differentiation,  
CC and/or function of cells that appear in the bone marrow, and leukocytes,  
CC treat bone marrow, blood and hematopoietic associated diseases and  
CC disorders, osteoarthritis, pulmonary congestion or oedema, emphysema,  
CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal  
CC disorders, spleen associated diseases, modulate renal disorders, treat  
CC cardiovascular disorders such as ischemic heart disease, modulate the  
CC proliferation, differentiation, and/or function of bone and cartilage  
CC cells and to treat bone and/or cartilage associated diseases or  
CC disorder. They may also be used to treat disorders associated with the  
CC ovaries, cerebral oedema, hydrocephalus, brain herniations, latrogenic  
CC disease, inflammations, bacterial and viral meningitis, Alzheimer's  
CC disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis  
CC brain cancers, hydrocephalus and encephalitis, and treat hepatic  
CC disorders.

Sequence 969 BP; 211 A; 316 C; 294 G; 148 T; 0 other;

Query Match	87.7%;	Score 755.2;	DB 21;	Length 969;
Best Local Similarity	99.6%;	Pred. No. 1.1e-145;		
Matches 757; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

0Y	102	GGAGATCAACGAGAGCTGGAGAGTTTGTTCACATATATCATCATCTGTGTGTGTGATGATGAT	167
Db	2	GGAGATGCGCGAGCGCTGGAGAGTTTGTTCACATATATCATCATCTGTGTGTGTGATGATGAT	167
0Y	162	GGTGTGTGTGATCAACGTGCTCTGTGAGCCATCTACAGTGTCTGCACGAGTCTTCATCAG	225
Db	62	GSTGTGTGTGATCAACGTGCTCTGTGAGCCATCTACAGTGTCTGCACGAGTCTTCATCAG	122
0Y	222	CCGGCACAACCCAGGGGGGGAGAGAGAAGATGCCCTGTCTCAGAAAGATGCCGTGTGGCC	281
Db	122	CCGGCACAACCCAGGGGGGGAGAGAGAAGATGCCCTGTCTCAGAAAGATGCCGTGTGGCC	181
0Y	282	CTCGAGAGAGCAACAGTGTAGAGCAACGGATATCCAGAGCCGAGAGCTTACGCCCGCCCTCG	342
Db	182	CTCGAGAGAGCAACAGTGTAGAGCAACGGATATCCAGAGCCGAGAGCTTACGCCCGCCCTCG	242
0Y	342	GCCCAACGACCGCCCTGGGCCGTTGCCGCTTCTTGCCACGAGGGAGCGCTTCCACCGCTTCCA	401
Db	242	GCCCAACGACCGCCCTGGGCCGTTGCCGCTTCTTGCCACGAGGGAGCGCTTCCACCGCTTCCA	301
0Y	402	GCCCAACCTATCCGTATCTCTGCACACACAGATGTGACCTGGCAACCCACATCTGCTGTGAGA	461
Db	302	GCCCAACCTATCCGTATCTCTGCACACAGATGTGACCTGGCCACACATCTGCTGTGAGA	361
0Y	462	CGGGAGAGAGCCCCCACCCTTACACAGGGGCCCTGTGCACCTTCAGAGTTGGGAGCCCCGAGCA	521
Db	362	CGGGAGAGAGCCCCCACCCTTACACAGGGGCCCTGTGCACCTTCAGAGTTGGGAGCCCCGAGCA	421
0Y	522	GCAGCTGGAAGTGAACCGGGAGGTGGGTGCGCGCACCCCCAAACAGAACATCTTTCAGACG	581
Db	422	GCAGCTGGAAGTGAACCGGGAGGTGGGTGCGCGCACCCCCAAACAGAACATCTTTCAGACG	481
0Y	582	TGACCTGATGATTAATGTGCAGAGGTGGGGGGCCCGCCCGCCACAGATTAATCTGGGGCAT	642
Db	482	TGACCTGATGATTAATGTGCAGAGGTGGGGGGCCCGCCCGCCACAGATTAATCTGGGGCAT	542
0Y	642	CAGCCCAACGTCTTACGGCAGAGGGGGCGATGAGAGGGGGCGCGCCACATTCACAGCA	701

Dd	542	CAGCGCCACGCTGTCTAACGGCAGCCGCCGGCGGCATGAAGGGCCCCGCCCACTTACAGCCA	601
Oy	702	GGTCTATCGGCACACTACC CGGGGTCCTCTTCCAGCACACAGAGAGTGCGCCCTC	761
Dd	602	GGTCTATCGGCACACTACC CGGGGTCCTCTTCCAGCACACAGAGAGTAGTGGCGCTC	661
Oy	762	CTTGTGTGAGGGGAGCCCGGCTCCACACACACACATCGCGCCCTTAAAGAGCGCAGCCAT	821
Dd	662	CTTGTGTGAGGGGAGCCCGGCTCCACACACACACATCGCGCCCTTAAAGAGCGCAGCCAT	721
Oy	822	CTGAGCAAAGAAGATTAACAAGAAAGACACCCCTC	861
Dd	722	CTGAGCAAAGAAGATTAACAAGAAAGACACCCCTC	761
<hr/>			
RESULT 3			
ID	ABK92120	standard; DNA; 1140 BP.	
XX	ABK92120;		
AC	15-AUG-2002	(first entry)	
DX			
DE	Prostate cancer-associated DNA sequence #6.		
XX			
KW	Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;		
KM	gene therapy; gene; ds.		
XX			
OS	Mammalia.		
XX			
PN	WO200230268-A2.		
PD	18-APR-2002.		
XX			
PF	12-OCT-2001; 2001MO-US32045.		
XX			
PR	13-OCT-2000; 2000US-0687576.		
PR	08-DEC-2000; 2000US-0733288.		
PR	08-DEC-2000; 2000US-0733742.		
PR	24-JAN-2001; 2001US-263957P.		
PR	16-MAR-2001; 2001US-276791P.		
PR	16-MAR-2001; 2001US-276888P.		
PR	06-APR-2001; 2001US-281922P.		
PR	24-APR-2001; 2001US-286214P.		
PR	30-APR-2001; 2001US-0847046.		
PR	04-MAY-2001; 2001US-288589P.		
XX			
PA	(EOSB-) EOS BIOTECHNOLOGY INC.		
XX			
Pi	Gish KC, Mack DH, Walson KE, Afar D, Hevezl P;		
XX			
DR	WPI: 2002-471335/50.		
XX	P-PSDB; ABG61805.		
XX			
PT	Detecting a prostate cancer-associated transcript in a cell in a		
PT	patient, useful for diagnosing prostate cancer (PC) or screening		
PT	modulators of PC, by determining if prostate cancer-associated genes		
PT	are expressed in a prostate tissue -		
XX			
XX			
PS	Claim 22; Page 305; 436pp; English.		
XX			
CC	The present invention relates to methods of detecting a prostate		
CC	cancer-associated transcript in a cell from a patient. The method		
CC	comprises contacting a biological sample from the patient with		
CC	prostate cancer-associated polynucleotides (designated PC genes) that		
CC	selectively hybridize to a sequence that is at least 80% identical		
CC	to them. The prostate cancer-associated polynucleotide sequences		
CC	are differentially expressed in prostate tumour tissue or in		
CC	prostate cancer and are derived from the tissues of various		
CC	organisms such as humans or other mammals (e.g. mice, sheep and dogs)		
CC	The methods of the invention are useful for diagnosing and treating		
CC	prostate cancer in mammals. The prostate cancer-associated genes are		
CC	useful for diagnosing or treating prostate cancer, as well as for		

CC Identifying modulators of prostate cancer or agents that inhibit  
 CC prostate cancer. The nucleic acid sequences are particularly useful  
 CC in gene therapy, as a vaccine or in antisense applications.  
 CC AKR92115-AKR92263 represent prostate cancer-associated polynucleotide  
 CC sequences.

XX Sequence 1140 BP; 270 A; 350 C; 336 G; 184 T; 0 other;  
 Query Match 87.6%; Score 754.2; DB 24; Length 1140;  
 Best Local Similarity 98.3%; Pred. No. 1,8e-145;  
 Matches 762; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

QY 87 TTTGTCAGAGCATGGAGATCAGGAGCTGGAGTTGTTCAGATCATCATCTGTCT 146
DB 77 TCTCTGGGAACCAAGCAATGGGGAGCTGGAGTTGTTCAGATCATCATCTGTCT 136
QY 147 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 146
DB 137 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 136
QY 207 ACGGTCCTTCATCAGGCGGCGACAGCCAGGGGGGAGAGAGAGAAAGTGGCCCTCTAGA 256
DB 197 ACGGTCCTTCATCAGGCGGCGACAGCCAGGGGGGAGAGAGAGAAAGTGGCCCTCTAGA 256
QY 267 AGGATGCTCTGTCGCTTCGAGAGCAGATGTCAGGGAGAGGAATCCAGAGCGGAGGT 326
DB 257 AGGATGCTCTGTCGCTTCGAGAGCAGATGTCAGGGAGAGGAATCCAGAGCGGAGGT 326
QY 327 CTAGCGCCCGCTCGGCGCCAGCCAGCGCTGGCCCTCGCCCTCGGCGAGGGAGCG 386
DB 317 CTAGCGCCCGCTCGGCGCCAGCCAGCGCTGGCCCTCGCCCTCGGCGAGGGAGCG 386
QY 387 CTTTCACGCGCTTCAGCCGACCTATCCGATCTGACAGAGAGATGACCTGACCCAG 446
DB 377 CTTTCACGCGCTTCAGCCGACCTATCCGATCTGACAGAGAGATGACCTGACCCAG 446
QY 447 CATCTGCTGTACAGCGGGAGAGGCGCCAGCTTACAGGGCCCTCTACCTCTGACCT 506
DB 437 CATCTGCTGTACAGCGGGAGAGGCGCCAGCTTACAGGGCCCTCTACCTCTGACCT 506
QY 507 TCGGGACCCCGAGCAGAGCTGGAAGTGAACCGGGAGTGGTGGCGGACCCCAACAG 566
DB 497 TCGGGACCCCGAGCAGAGCTGGAAGTGAACCGGGAGTGGTGGCGGACCCCAACAG 566
QY 567 AACCATCTTCGACAGTGCATGATGATGATGATGATGATGATGATGATGATGATGAT 626
DB 557 AACCATCTTCGACAGTGCATGATGATGATGATGATGATGATGATGATGATGATGAT 626
QY 627 CAGTAACTGGGGCATCAGCGGCGAGTGTACAGCGGGGGCGCATGGAGGGGCGGCC 686
DB 617 CAGTAACTGGGGCATCAGCGGCGAGTGTACAGCGGGGGCGCATGGAGGGGCGGCC 686
QY 687 GCCCACTACAGCGAGGTCATGCGGCACTACCCGGGGTCTCTCTTCACACACCAAG 746
DB 677 GCCCACTACAGCGAGGTCATGCGGCACTACCCGGGGTCTCTCTTCACACACCAAG 746
QY 747 CAGTGGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 806
DB 737 CAGTGGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 806
QY 807 AGAGAGCGCAGCATCTGGAGCAAGAGATTAACAGAAAGAGACCTCTCTC 861
DB 797 AGAGAGCGCAGCATCTGGAGCAAGAGATTAACAGAAAGAGACCTCTCTC 861

```

RESULT 4  
 ID AAA75163 standard; cDNA; 969 BP.  
 AC AAA75163;  
 XX  
 DT 15-JAN-2001 (first entry)  
 XX

DE cDNA clone encoding a human TANGO 261 polypeptide.

XX TANGO 266; TANGO 216; TANGO 261; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW edema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; latrogenic disease; inflammation; meningitis;  
 KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FT CDS 6..764  
 FT /tag= a  
 FT /product= "TANGO 261"

W0200052022-A1.

08-SEP-2000.

01-MAR-2000; 2000MO-US05226.

01-MAR-1999; 9905-0122458.

(MILL-) MILLENNIUM PHARM INC.

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

DR WPI; 2000-579269/54.

DR P-PSDB; AAB18461.

PT Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -

PS Disclosure; Page -: 175pp; English.

CC AAA75163-65 encode human TANGO 261 proteins. The specification also  
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat any von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
 CC and cell trafficking and/or migration, modulate cellular interactions,  
 CC modulate cell adhesion in proliferative disorders, such as cancer,  
 CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 CC and hematopoietic associated diseases and disorders, treat bone marrow, blood  
 CC pulmonary congestion or edema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, spleen associated  
 CC diseases, modulate renal disorders, treat cardiovascular disorders such  
 CC as ischemic heart disease, modulate the proliferation, differentiation,  
 CC and/or function of bone and cartilage cells and to treat bone and/or  
 CC cartilage associated diseases or disorder. They may also be used to  
 CC treat disorders associated with the ovaries, and cerebral edema,  
 CC hydrocephalus, brain herniations, latrogenic disease, inflammation,  
 CC bacterial and viral meningitis, Alzheimer's disease, cerebral  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC hydrocephalus and encephalitis, and treat hepatic disorders.  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided.

XX Sequence 969 BP; 211 A; 317 C; 293 G; 148 T; 0 other;

Query Match 87.5%; Score 753.6; DB 21; Length 969;  
 Best Local Similarity 99.5%; Pred. No. 2.3e-145;  
 Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 102 GGAGTACAGGAGCTGGAGTTGTTCAGATCATCATCTGTGTGTGTGTGTGTGTGT 161



Db 182 CTGGAGAGACAGTGTGACGACGAGATCCACAGAGCCGAGTCTACGCCGCCCTCG 241  
 QY 342 GCCCACCAGACCGCTGGCCGCTGCCCTTGGCCAGCGGAGGCTTCCACCGCTTCA 401  
 Db 242 GCCCACCAGACCGCTGGCCGCTGCCCTTGGCCAGCGGAGGCTTCCACCGCTTCA 401  
 QY 402 GCCCACCAGACCGCTGGCCGCTGCCCTTGGCCAGCGGAGGCTTCCACCGCTTCA 301  
 Db 302 GCCCACCAGACCGCTGGCCGCTGCCCTTGGCCAGCGGAGGCTTCCACCGCTTCA 461  
 QY 462 CGGGAGAGAGCCGCCACCCCTACAGGCGCCCTTGCACCTTCCAGGACCCGAGCA 521  
 Db 362 CGGGAGAGAGCCGCCACCCCTACAGGCGCCCTTGCACCTTCCAGGACCCGAGCA 421  
 QY 522 GCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCCAACACATCTTCTGACAG 581  
 Db 422 GCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCCAACACATCTTCTGACAG 481  
 QY 582 TGACCTGATGATAGTGGCGGAGTGGGCGCCCTTGCACCTTCCAGGACCCGAGCA 641  
 Db 482 TGACCTGATGATAGTGGCGGAGTGGGCGCCCTTGCACCTTCCAGGACCCGAGCA 541  
 QY 642 CACGCGCAGTGTCTACGCGAGCGGCGGCGATGGAGGGGCGCGCCGCTTACGCGAT 701  
 Db 542 CACGCGCAGTGTCTACGCGAGCGGCGGCGATGGAGGGGCGCGCCGCTTACGCGAT 601  
 QY 702 GGTGATGCGCCACTACCCGCGGCTCTCTTCCAGCAGCAGAGAGTGGCGCGCTT 761  
 Db 602 GGTGATGCGCCACTACCCGCGGCTCTCTTCCAGCAGCAGAGAGTGGCGCGCTT 661  
 QY 762 CTTCGCGAGGAGGCGCGGCTCTCTTCCAGCAGCAGAGAGTGGCGCGCTT 821  
 Db 662 CTTCGCGAGGAGGCGCGGCTCTCTTCCAGCAGCAGAGAGTGGCGCGCTT 721  
 QY 822 CTTCGCGAGGAGGCGCGGCTCTCTTCCAGCAGCAGAGAGTGGCGCGCTT 861  
 Db 722 CTTCGCGAGGAGGCGCGGCTCTCTTCCAGCAGCAGAGAGTGGCGCGCTT 761

RESULT 6  
 AAA75165  
 ID AAA75165 standard; cDNA; 969 BP.  
 XX  
 AC AAA75165;  
 XX  
 DT 15-JAN-2001 (first entry)  
 XX  
 DE cDNA clone encoding a human TANGO 261 polypeptide.  
 XX  
 KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hemotopoietic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 6...764  
 FT CDS /tag- a  
 XX /product- "TANGO 261"  
 XX  
 PD WO200052022-A1.  
 XX  
 PD 08-SEP-2000.  
 XX  
 PF 01-MAR-2000; 2000WO-US05226.

PR 01-MAR-1999; 99US-0122458.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
 XX  
 DR WPI; 2000-579269/54.  
 DR P-PSDB; AAB18463.  
 XX  
 PT Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PS e.g. for treating cancer -  
 PS  
 PS Disclosure; Page -: 175pp; English.

AAA75163-65 encode human TANGO 261 proteins. The specification also  
 describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 polypeptides can be used to modulate cellular proliferation, modulate  
 cellular differentiation and/or modulate cellular adhesion. The  
 proteins can be used to treat any von Willebrand factor-associated  
 disorder, regulate extracellular matrix structuring, cellular adhesion  
 and cell trafficking and/or migration, modulate cellular interactions,  
 modulate cell adhesion in proliferative disorders, such as cancer,  
 that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 and hematopoietic associated diseases and disorders, atelectasis, blood  
 pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
 diseases, modulate renal disorders, treat cardiovascular disorders such  
 as ischemic heart disease, modulate the proliferation, differentiation,  
 and/or function of bone and cartilage cells and to treat bone and/or  
 cartilage associated diseases or disorder. They may also be used to  
 treat disorders associated with the ovaries, and cerebral oedema,  
 hydrocephalus, brain herniations, iatrogenic disease, inflammation,  
 bacterial and viral meningitis, Alzheimer's disease, cerebral  
 toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 note: the present sequence does not appear in the specification. It was  
 created using information provided.

Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other;  
 SO

Query Match 87.5%; Score 753.6; DB 21; Length 969;  
 Best Local Similarity 99.5%; Pred. No. 2.3e-145;  
 Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 102 GGAGATCAGGAGCTGGAGTGTGTCAGATCATCATCGTGGGATGATGATGAT 161  
 Db 2 GGAGATGCGGAGCTGGAGTGTGTCAGATCATCATCGTGGGATGATGATGAT 61  
 QY 162 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 221  
 Db 62 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121  
 QY 222 CCGGACAGCGAGGCGGAGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 281  
 Db 122 CCGGACAGCGAGGCGGAGGAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 181  
 QY 282 CTGGAGAGACAGTGTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 341  
 Db 182 CTGGAGAGACAGTGTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 241  
 QY 342 GCCCACCAGACCGCTGGCCGCTGCCCTTGGCCAGCGGAGGCTTCCACCGCTTCA 401  
 Db 242 GCCCACCAGACCGCTGGCCGCTGCCCTTGGCCAGCGGAGGCTTCCACCGCTTCA 301  
 QY 402 GCCCACCAGACCGCTGGCCGCTGCCCTTGGCCAGCGGAGGCTTCCACCGCTTCA 461  
 Db 302 GCCCACCAGACCGCTGGCCGCTGCCCTTGGCCAGCGGAGGCTTCCACCGCTTCA 361  
 QY 462 CGGGAGAGAGCCGCCACCCCTACAGGCGCCCTTGCACCTTCCAGGACCCGAGCA 521  
 Db 362 CGGGAGAGAGCCGCCACCCCTACAGGCGCCCTTGCACCTTCCAGGACCCGAGCA 421

Oy	522	GCAGCTGGAACTGAAACCGGGGAGTGTGGTCGGGCGACACCCCAAAACAGAACCACTTCGCACAG	581
.Db	422	GCAGCTGGAACTGAAACCGGGGAGTGTGGTCGGGCGACACCCCAAAACAGAACCACTTCGCACAG	481
Oy	582	TGACCTGATGATAGTAGTGCACGAGCTGGGGGGGCCCTGGCCCCCCCCAGCAGTAATCTGGGGCAT	641
Db	482	TGACCTGATGATAGTAGTGCACGAGCTGGGGGGGCCCTGGCCCCCCCCAGCAGTAATCTGGGGCAT	541
Oy	642	CAGCGCCACAGTGTCTACGCGCAGCGCGGGCGCATGAGGGGGCGCGCCACACTACAGCGA	701
Db	542	CAGCGCCACAGTGTCTACGCGCAGCGGGCGCATGAGGGGGCGCGCCACACTACAGCGA	601
Oy	702	GGTCTATGGGCCACTACCCGGGGGTCTCTCTTCACAGCACCAGCAGAGCATGGGGCGCCCTC	761
Db	602	GGTCTATGGGCCACTACCCGGGGGTCTCTCTTCACAGCACCAGCAGAGCATGGGGCGCCCTC	661
Oy	762	CTTGCTGGAGGGGAGACCGGGCTCCACACACACACATCGCGCCCTTAGAGAGCGACGCAT	821
Db	662	CTTGCTGGAGGGGAGACCGGGCTCCACACACACACATCGCGCCCTTAGAGAGCGACGCAT	721
Oy	822	CTGAGCAAGAGGAAGATTAACAGAAAGGACACCCCTGC	881
Db	722	CTGAGCAAGAGGAAGATTAACAGAAAGGACACCCCTGC	781

RESULT 7	
AAAA7429	
ID	AAAA7429 standard; DNA; 1061 BP.
XX	
AC	AAAA7429;
XX	
DT	20-OCT-2000 (first entry)

DE Sequence encoding human neuron-associated protein

KW Neuron associated protein; NEAP; neurological disorder; epilepsy;  
KW ischaemic cerebrovascular disease; stroke; cerebral neoplasm;  
KW Alzheimer's disease; Pick's disease; Huntington's disease;  
KW dementia; Parkinson's disease; demyelinating disease; meningitis;  
KW prion disease; Kurru; Creutzfeldt-Jakob disease; neurofibromatosis;  
KW cerebral palsy; muscular dystrophy; central nervous system;  
KW peripheral nervous system; PNS; myopathy; schizophrenia;  
KW acral keratosis; arteriosclerosis; atherosclerosis; buritis;  
KW cirrhosis; hepatitis; mixed connective tissue disease; MCMD;  
KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
KW AIDS; Addison's disease; adult respiratory distress syndrome;  
KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
KW Werner syndrome; trauma; human; ds

OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	101..859

FT	/*tag= a
FT	/product= Neuron associated protein

PN WO200034477-A2

PD 15-JUN-2000.

PF 10-DEC-1999; 99WO-US30408.

PR 11-DEC-1998; 98US-0210083.

PR 09-FEB-1999; 99US-0119365.

XX

XX XX

PI LU DAM, AZIMZAI Y;

XX WP1: 2000-423423/36.  
 DX  
 XX P-PSDB: AAB01388.  
 DX  
 XX  
 PT New human neuron-associated proteins and polynucleotides encoding them  
 PT useful for diagnosis, treatment and prevention of cell proliferative  
 PT disorders including cancer, neuronal and neurological disorders  
 XX  
 PS Claim 9; Page 136; 145pp; English.

Human neuron-associated proteins (NEUP) can be used for for treating or preventing a disorder associated with decreased expression or activity of NEUP. Antagonists of NEUP are useful for treating or preventing disorder associated with increased expression or activity of NEUP. NEUP or their fragments or derivatives are useful for treating neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia and Parkinson's disease. NEUPs are also useful for treating other demyelinating diseases, bacterial and viral meningitis, prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and metabolic diseases of the nervous system, neurofibromatosis, other developmental disorders of the central nervous system, cerebral palsy, neuromuscular disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, inherited, metabolic, endocrine, and toxic myopathies, mental disorders including mood, anxiety and schizophrenic disorders, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal haemoglobinuria, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. This sequence was given the InCyte ID no. 1871288CB1.

Sequence 1061 BP; 225 A; 342 C; 326 G; 168 T; 0 other;

Query Match 87.48; Score 752.6; DB 21; Length 1061;

```
Matches 761; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

87 TTTGTTCCAGAGCATGGAGATCACGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT 146

Db 82 TCTCTGGAACGAGCAATGGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGT 141

QY 147 GGTGATGGTGGTGATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGC 206

Db 142 GGTGATGATGGTGGTGATCAGCTGCCCTGCTGAGCCACTACAAGCTGTCTGC 201

207 ACGGTCCTTCATCAGCCGGCACAGCCAGGGGGCGAGAGAGAGATGCCCTGTCCTCAGA 266

Db 202 ACGGTCCTTCATCAGCCGGCACAGCCAGGGGGCGAGAGAGAGATGCCCTGTCTCAGA 261

267 AGGATGCCCTGTGGCCCTCGGAGAGCAGTGTCAGGCAACGGAATCCAGAGCCGCAGGT 326

Db 262 AGGATGCCCTGTGCCCTCGGAGAGCACAGTGTcAGGCAACGGAATCCcAGAGCCGcAGGT 321

QY 327 CTACGCCCCGCTCGGCCACCGACCGCTGGCCGTGCCGCCCTTCGCCACGGGAGCG 386

Db 322 CTACGCCCCGCTCGGCCACCGACCGCTGGCCGTGCCGCCCTTCGCCACGGGAGCG 381

387 CTTCACCGCTTCAGCCACCTATCCGTACTGCAGCAGAGATCGACCTGCCACCCAC 446

Db 382 CTTCCACCGCTTCAGCCACCTATCCGTACTGCAGCAGAGATCGACCTGCCGCCAC 441

447 CATCTCGCTGTACAGACGGGGAGGAGCCCCACCCCTACCAAGGGCCCTGCACCCCTCAGCT 506

Sequence 1000 BP; 225 A; 348 C; 325 G; 168 T; 0 other;

Query Match	86.2%;	Score 742.4;	DB 22;	Length 1066;
Best Local Similarity	99.2%;	Pred. No. 4.6e-143;		
Matches 746; Conservative	0.0	Mismatches		

Matches	Conservative	Mismatches
746	0	6

CGGAGCTGGAGCTTTCACCA

Db 113 CGACGACGATCTTCTTCAGCAGTGGTGCG 16

.....CAGATCATTCATCGTGGTGGTATGATGCTGATGGTGGTGG 17

170 TGATCAGCTGCCCTGCTGAGCCACTACAGAATTCTTTCAGCGGGG

pb  
173 mcaatccttccacggcgaca 22

175 181CAGTGGCTGAGCCACTACAAGCTGTCTGCAGGTCCTTATCAGCCCAAG 220

[illegible][illegible]

D<sub>b</sub> 233 GCCAGGGCGGAGGAGAATGCCCTGTCTTCAACACCATCCCTA

-----  
 ..ccacnaggaagccctgaggaga 292

250 GCACAGTGTTCAGGCAACGGATCCAGAGCCCGCAGGTCTACGCCCGCTTCGAGCCGACCG

[illegible]

...GCGGAGGCGGAGGCTACGCCCCGCTCGGCCACCG 352

350 ACCGCGTGGCCGTGCCGCCCTTCGCAAGCACAACATCATTA

|||||CCACCGCTTCCAGCCACCT 409

555 AUCGCUUGCCGTCGCCCTTCGCCCAGCGGAGCGCTTCCACCGTTCAACGCGCGCGT

.....

410 ATCCGACCTCAGCGGGTTC  
..... 412

[illegible]

413 ATCCGTACCTGCAGCAGGATGACTTCCTACCGA

.....CCACCAATCTCGCTGTACAGACGGGAGG 472

**4**/0 AGCCCCCACCCTACCAAGGGCCCTGCACCCTCACGTTGGGAACCGAACAAGAATTA

b  
473 AGCCCGGCTGGTGTCTTTATAGCTCCTGACCTGCTGG 529

.....  
...cccccccaacccaatcgcacccttcagcttccgggacctccgaggagcacgtcg 537

Y 530 AACTGACCGGGAGTCGGTGCCTCCCAACCCCAAAATTT

		cccccAAACAGAACCATCTCGACAGTGACCCTGA 589	

533 AACTGACCGGGAGTCGTGCCGCACCCCAACACAGATCTTCCACCTC

TGCAATACCTCCCGAAGGGTGTTCGTAAGTAGCCTTGGA 592

1000 5916CCAGGCTGGGCGGCCCTGCCCGCCCGCAGCAGTAACTCGGCATCAGGCCA 649

593 TGGATAGTGCACGCTGGGCCCGCCCTCCTCCGCTCC

.....GCCCCCGCCCCGAGCAGTACTCGGGCATCAGGCCA 652

650 CGTGCTACGGCAGCGGCGGCGCATGGAGGCGCCGCCGC

|||||CCCCCCTTACAGCGAGGTTCATCG 709  
653 CCGCCGTTGGT

555 CAGTACGGCAGCGGCGGCATGGAGGGGCGCGCCGCCACCTACAGCGAGTATTG 715

**710** GCCACTATCGGCTACATTGGAGTGTGGTGTTGGGTCGGCCTCC  
.....GCAACAGGCAG / 112

**CCAGCAGCAGTGGGCCGCTCCTTGCTGG**

713 GCCACTACCGGGGTCCTCCTCCAGCACCAAGAACGCATCTGCGGTTT

770 .....  
772

170 AGGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGTACGCAATTGG

773 AGGGACCCCCCTGCA 829





[illegible]

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
 XX  
 DR WPI: 2000-579269/54.  
 DR P-PSDB: AAB18450.

XX Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 261, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -

XX Claim 2; Fig 6A-B; 175pp; English.

XX The present sequence encodes a murine TANGO 261 polypeptide. The  
 CC specification also describes TANGO 266, TANGO 216, TANGO 267, and  
 CC TANGO 267. The TANGO polypeptides can be used to modulate cellular  
 CC proliferation, modulate cellular differentiation and/or modulate  
 CC cellular adhesion. The proteins can be used to treat any von Willebrand  
 CC factor-associated disorder, regulate extracellular matrix structuring,  
 CC cellular adhesion, and cell trafficking and/or migration, modulate  
 CC cellular interactions, modulate cell adhesion in proliferative  
 CC disorders, such as cancer, modulate the proliferation, differentiation,  
 CC and/or function of cells that appear in the bone marrow, and leukocytes,  
 CC treat bone marrow, blood and hematopoietic associated diseases and  
 CC disorders, atelectasis, pulmonary congestion or edema, emphysema,  
 CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal  
 CC disorders, spleen associated diseases, modulate renal disorders, treat  
 CC cardiovascular disorders such as ischemic heart disease, modulate the  
 CC proliferation, differentiation, and/or function of bone and cartilage  
 CC cells and to treat bone and/or cartilage associated diseases or  
 CC disorder. They may also be used to treat disorders associated with the  
 CC ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic  
 CC disease, inflammations, bacterial and viral meningitis, Alzheimer's  
 CC disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,  
 CC brain cancers, hydrocephalus and encephalitis, and treat hepatic  
 CC disorders.

XX Sequence 1713 BP; 506 A; 438 C; 400 G; 369 T; 0 other;

XX Query Match 53.7%; Score 462.2; DB 21; Length 1713;

XX Best Local Similarity 83.8%; Pred. No. 1.2e-85; Mismatches 27; Gaps 3;

XX Matches 569; Conservative 0; Mismatches 83; Indels 27; Gaps 3;

QY 183 GCTGAGCCATACAACTGTCTGACGCTCTTCATACAGCCGACAGGCGGAG 242  
 Db 1 GCTGAGCCATACAACTGTCTGACGCTCTTCATACAGCCGACAGGCGGAG 60  
 QY 243 GAGAGAGATGCTCTGTCTTCAAGAAAGATGCTGTGCTTGGACACAGTGTCAAG 302  
 Db 61 GAGAGAGATGCTCTGTCTTCAAGAAAGATGCTGTGCTTGGACACAGTGTCAAG 120  
 QY 303 CAAGGAAATCCAGAGCGGAGGCTTACGCGCCGCTGGGCGCACAGCGCTGGCGGT 362  
 Db 121 ---TGAATGCGCGAGGACAGGTATGCCCCGCTGGCCCATGACCACTGCTGCT 177  
 QY 363 GCGGCGCTTTCGCGGAGCGGAGCGCTTCCAGCGCTTCCAGCTATCCGTACTCTCA 422  
 Db 178 GCGGCGCTTTCGCGGAGCGGAGCGCTTCCAGCGCTTCCAGCTATCCGTACTCTCA 228  
 QY 423 GCAGAGATGCACTGCTGCAACCACTATCGCTGTCAAGGGGAGAGCCGCCCTTA 482  
 Db 229 GCAGAGATGCACTGCTGCAACCACTATCGCTGTCAAGGGGAGAGCCGCCCTTA 288  
 QY 483 CCAGGCGCCCTGCAACCTTCCAGCTTCCAGCGGAGGAGAGAGCACTGAACCGGGA 542  
 Db 289 CCAGGCGCCCTGCAACCTTCCAGCTTCCAGCGGAGGAGAGAGCACTGAACCGGGA 348  
 QY 543 GTGCGTGGCGACCCCAACAGAACCATCTTTCAGACAGTCACTGATGATAGTCCAG 602  
 Db 349 ATGCTGGGCGAGCCCTTAACCGGACCATCTTTCAGACAGTCACTGATGATAGTCCAG 408  
 QY 603 GCTGGGCGGCGCCCTTCCAGCACTGATGATGATGATGATGATGATGATGATGAT 662  
 Db 409 GCTGGGCGGCGCCCTTCCAGCACTGATGATGATGATGATGATGATGATGATGAT 468

OY 663 CGGGGCGGCGATGAGAGGCGCGCCGACCTACAGGAGGTCATCGGCCCATACCGGG 722  
 Db 469 CGGTGGGCGATGAGAGGCGCGCCGACCTACAGGAGGTCATCGGCCCATACCGGG 528  
 OY 723 GTCTCTTTCAGAGCAACGAGAGAGGAGGCGCGCTCTCTTCTGTGAGGAGCCGCT 782  
 Db 529 CTCCTCTTTCAGAGCAACGAGAGAGGAGGCGCGCTCTCTTCTGTGAGGAGCCGCT 588  
 OY 783 CCACCACACACATCGCGCCCTTACAGAGCGGAGGAGGAGGAGGAGGAGGAGG 842  
 Db 589 CCATCTCTGAGCACTTGGCCGACCTGGA-----GAACAAGAGAGAGAGGAA 633  
 OY 843 ACAGAAAGAGACCTCTTC 861  
 Db 634 ACAGAAAGTCAACCCCTTC 652

RESULT 12  
 AAA75166

ID AAA75166 standard; cDNA: 1713 BP.

XX AAA75166;

DT 15-JAN-2001 (first entry)

DE cDNA clone encoding a murine TANGO 261 polypeptide.

XX TANGO 266; TANGO 216; TANGO 261; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW edema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

XX Mus sp.

XX Key Location/Qualifiers

FT CDS 2..655

FT /tag= a

FT /product= "TANGO 261"

XX WO200052022-A1.

XX 08-SEP-2000.

XX 01-MAR-2000; 2000WO-US05226.

XX 01-MAR-1999; 99US-0122458.

XX (MILL-) MILLENNIUM PHARM INC.

XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

XX WPI: 2000-579269/54.

XX P-PSDB: AAB18464.

XX Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 261, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -

XX Disclosure: Page -: 175pp; English.

XX AAA75166-68 encode murine TANGO 261 proteins. The specification also  
 CC describes TANGO 266, TANGO 216, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat any von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
 CC and cell trafficking and/or migration, modulate cellular interactions,



XX Sequence 1713 BP; 506 A; 439 C; 399 G; 369 T; 0 other;

Query Match 53.5%; Score 460.6; DB 21; Length 1713;

Best Local Similarity 83.7%; Pred. No. 2.6e-85; Matches 568; Conservative 0; Mismatches 84; Indels 27; Gaps 3;

```

OY 183 GCTAGGCACTACAAAGCTGCTCCTCATCAGCCGCGACAGCCAGGCGGAG 242
DB 1 GCTAGGCACTACAAAGCTGCTCCTCATCAGCCGCGACAGCCAGGCGGAG 60
OY 243 GAGAGAGATGCTCTCTCAGAGAGATGCTGCGCTCGAGAGACAGTGTACG 302
DB 61 GAGAGAGATGCTCTCTCAGAGAGATGCTGCGCTCGAGAGATGCTGAGG 120
OY 303 CAACGGAATCCCAAGCGGAGGCTTACGCGCCGCTGCGCCACCGAGCGCTG 362
DB 121 ---TGGATGCGGAGGAGGCTTATGCGCCCTGCGCCACCTGACCTGCTGT 177
OY 363 GCGGCTTGGCGGAGGCGGAGGCTTCCACGCTGCGCCACCTGACCTGCA 422
DB 178 GCGGCTTGGCGGAGGCGGAGGCTTCCACGCTGCGCCACCTGACCTGCA 228
OY 423 GCACGAGATGACCTGCGCACCCACCATCTGCTGACAGCGGAGGAGCCACCTA 482
DB 229 GCACGAGATGACCTGCGCACCCACCATCTGCTGACAGCGGAGGAGCCACCTA 288
OY 483 CCAGGCGGCTGACCTGCGCACCTGCGGAGCGGAGGAGGAGGAGGAGGAG 542
DB 289 CCAGGCGGCTGACCTGCGCACCTGCGGAGCGGAGGAGGAGGAGGAGGAG 348
OY 543 GTGGTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602
DB 349 ATCTGTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 408
OY 603 GCTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 662
DB 409 GCTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 468
OY 663 CGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
DB 469 CGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 528
OY 723 GTCTCTCTCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
DB 529 CTCTCTCTCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588
OY 783 CCACGAGACACATCTGCGCGGCTGAGAGGCGGAGGAGGAGGAGGAGGAG 842
DB 589 CCATCACTCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 633
OY 843 ACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
DB 634 ACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 652

```

RESULT 14

AAA75168  
ID AAA75168 standard; cDNA; 1713 BP.

XX AAA75168;

XX 15-JAN-2001 (first entry)

XX cDNA clone encoding a murine TANGO 261 polypeptide.

XX TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;

KW cellular proliferation; cellular differentiation; cellular adhesion;

KW von Willebrand factor-associated disorder; cell trafficking; cancer;

KW hematopoietic associated disease; atelectasis; pulmonary congestion;

KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;

KW intestinal disorder; spleen associated disease; renal disorder;

KW cardiovascular disorder; ischemic heart disease; hydrocephalus;

KW brain herniation; iatrogenic disease; inflammation; meningitis;

KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;

KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

OS Mus sp.

FT Key Location/Qualifiers

FT CDS 2..655

FT /tag= a

FT /product= "TANGO 261"

XX W0200052022-A1.

XX 08-SEP-2000.

XX 01-MAR-2000; 2000WO-US05226.

XX 01-MAR-1999; 99US-0122458.

XX (MILL-) MILLENNIUM PHARM INC.

XX Barnes TM, Holzman DA, Sharp JD, Fraser CC;

XX WPI: 2000-579269/54.

XX P-PSDB: AAB18466.

XX Novel human and murine secreted proteins designated TANGO 216, 261,

XX 262, 266 and 267 useful as modulating agents of cellular processes,

XX e.g. for treating cancer -

XX Disclosure; Page -: 175pp; English.

XX AAA75166-68 encode murine TANGO 261 proteins. The specification also

XX describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO

XX polypeptides can be used to modulate cellular proliferation, modulate

XX cellular differentiation and/or modulate cellular adhesion. The

XX proteins can be used to treat any von Willebrand factor-associated

XX disorder, regulate extracellular matrix structuring, cellular adhesion,

XX and cell trafficking and/or migration, modulate cellular interactions,

XX modulate cell adhesion in proliferative disorders, such as cancer,

XX modulate the proliferation, differentiation, and/or function of cells

XX that appear in the bone marrow, and leukocytes, treat bone marrow, blood

XX and hematopoietic associated diseases and disorders, atelectasis,

XX pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial

XX asthma and bronchiectasis, intestinal disorders, spleen associated

XX diseases, modulate renal disorders, treat cardiovascular disorders such

XX as ischemic heart disease, modulate the proliferation, differentiation,

XX and/or function of bone and cartilage cells and to treat bone and/or

XX cartilage associated diseases or disorder. They may also be used to

XX treat disorders associated with the ovaries, and cerebral oedema,

XX hydrocephalus, brain herniations, iatrogenic disease, inflammations,

XX bacterial and viral meningitis, Alzheimer's disease, cerebral

XX toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,

XX hydrocephalus, and encephalitis, and treat hepatic disorders.

XX note: the present sequence does not appear in the specification; it was

XX created using information provided.

XX Sequence 1713 BP; 506 A; 439 C; 399 G; 369 T; 0 other;

XX AAA75168

XX Query Match 53.5%; Score 460.6; DB 21; Length 1713;

XX Best Local Similarity 83.7%; Pred. No. 2.6e-85;

XX Matches 568; Conservative 0; Mismatches 84; Indels 27; Gaps 3;

```

OY 183 GCTAGGCACTACAAAGCTGCTCCTCATCAGCCGCGACAGCCAGGCGGAG 242
DB 1 GCTAGGCACTACAAAGCTGCTCCTCATCAGCCGCGACAGCCAGGCGGAG 60
OY 243 GAGAGAGATGCTCTCTCAGAGAGATGCTGCGCTCGAGAGACAGTGTACG 302
DB 61 GAGAGAGATGCTCTCTCAGAGAGATGCTGCGCTCGAGAGATGCTGAGG 120
OY 303 CAACGGAATCCCAAGCGGAGGCTTACGCGCCGCTGCGCCACCGAGCGCTG 362
DB 121 ---TGGATGCGGAGGAGGCTTATGCGCCCTGCGCCACCTGACCTGCTGT 177
OY 363 GCGGCTTGGCGGAGGCGGAGGCTTCCACGCTGCGCCACCTGACCTGCA 422
DB 178 GCGGCTTGGCGGAGGCGGAGGCTTCCACGCTGCGCCACCTGACCTGCA 228
OY 423 GCACGAGATGACCTGCGCACCCACCATCTGCTGACAGCGGAGGAGCCACCTA 482
DB 229 GCACGAGATGACCTGCGCACCCACCATCTGCTGACAGCGGAGGAGCCACCTA 288
OY 483 CCAGGCGGCTGACCTGCGCACCTGCGGAGCGGAGGAGGAGGAGGAGGAG 542
DB 289 CCAGGCGGCTGACCTGCGCACCTGCGGAGCGGAGGAGGAGGAGGAGGAG 348
OY 543 GTGGTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602
DB 349 ATCTGTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 408
OY 603 GCTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 662
DB 409 GCTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 468
OY 663 CGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
DB 469 CGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 528
OY 723 GTCTCTCTCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
DB 529 CTCTCTCTCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588
OY 783 CCACGAGACACATCTGCGCGGCTGAGAGGCGGAGGAGGAGGAGGAGGAG 842
DB 589 CCATCACTCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 633
OY 843 ACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
DB 634 ACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 652

```

Db 121 ---TGGAATGCCGACACACAGTCTAATGCCCCGCTCGGCCACCTGATACCGACTCGCTGT 177  
 QY 363 GCCGCCCTTGGCCAGGGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCA 422  
 Db 178 GCCGCCCTTCAATCCAGCGG-----AGCCGATTCACCAACCCACCTACCCCTACTGCA 228  
 QY 423 GCACGAGATGCACTGGCCACCAACATCTGCTGTGAGAGGGGAGAGAGCCGCCACTTA 482  
 Db 229 GCACGAAATGCGCTGGCCACCAACATCTGCTGTGATGGGAGAGAGCCGCCACTTA 288  
 QY 483 CCAGGCGCCCTGCAACCTTCCAGCTTGGGAGCCCGAGACAGCACTGGAATGAACCGGGA 542  
 Db 289 CCAGGCGCCCTGCAACCTTCCAGCTTGGGAGCCCGAGACAGCACTGGAATGAACCGGGA 348  
 QY 543 GTGGGTGCGGACACCCCAAGAACCATCTTGCAGAGTGAACCTGATGATGTCAG 602  
 Db 349 ATCTGTGCGGACACCCCTTACCGGACCATCTTGCAGATGATGATGATGATGTCAG 408  
 QY 603 GCTGGGCGGCGCCCTGCGCCCGCCAGCACTGAGTGGGATGAGCGGACGCTGACGGGAG 662  
 Db 409 GCTGGGCGGCGCCCTGCGCCCGCCAGCACTGAGTGGGATGAGCGGACGCTGACGGGAG 468  
 QY 663 CGGCGGCGGCGGAGGCGCGCCCGCCAGCACTGAGTGGGATGAGCGGACGCTGACGGGAG 722  
 Db 469 CGGTCGCGGCGGAGGAGGCGCGCCCGCCAGCACTGAGTGGGATGAGCGGACGCTGACGGGAG 528  
 QY 723 GTCTCTCTTCCAGCAACCAAGCACTGAGTGGGATGAGCGGACGCTGACGGGAG 782  
 Db 529 GTCTCTCTTCCAGCAACCAAGCACTGAGTGGGATGAGCGGACGCTGACGGGAG 588  
 QY 783 CCACGACACACACATGCGGCGCCCTGAGAGAGCGGACGCTGAGTGGGATGAGCGGAG 842  
 Db 589 CCATGACTGCGACATGCGGCGCCCTGAGAGAGCGGACGCTGAGTGGGATGAGCGGAG 633  
 QY 843 ACAGAAAGGACACCCCTCTC 861  
 Db 634 ACAGAAAGGACACCCCTCTC 652

RESULT 15  
 ABR12143/C  
 ID ABR12143 standard; cDNA; 693 BP.  
 AC ABR12143;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE Human MIVR-1 homologous sequence #1.  
 XX  
 KW Human; ss; MIVR-1; Mechanically Induced Vascular Receptor 1;  
 KW cytoskeletal; cardiant; cerebroprotective; antiarteriosclerotic;  
 KW cardiac cell; anti-apoptotic; vascular endothelial cell;  
 KW heart failure; myocardial infarction; stroke; arteriosclerosis;  
 KW heart failure; A1761441.1.  
 XX  
 OS Homo sapiens.  
 PN W0200216416-A2.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 21-AUG-2001; 2001WO-US26089.  
 XX  
 PR 22-AUG-2000; 2000US-227159P.  
 XX  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
 PA (PFIZ ) PFIZER INC.  
 PI Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;  
 XX WPI; 2002-280912/32.  
 XX  
 PT Novel nucleic acid molecule encoding Mechanically Induced Vascular

PT Receptor-1 polypeptide, useful for treating cardiovascular diseases  
 XX  
 PS Disclosure; Page 101; 105pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having  
 CC cardiac cell anti-apoptotic activity and fragments of it provided  
 CC they are not identical to Genbank sequences A1761441.1, A1594390,  
 CC NM\_004338 and A0177461. Also included are expression vectors, host  
 CC cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of  
 CC MIVR-1, contacting a molecule having cardiac cell anti-apoptotic activity  
 CC with a candidate agent, where the molecule is a nucleic acid molecule  
 CC comprising MIVR-1, IEX-1, VDR-1, BRG-2 and TIS-110 or its  
 CC expression product, determining if the anti-apoptotic activity is  
 CC modulated and thereby identifying a modulator. The cardiac cell anti-  
 CC apoptotic molecules and nucleic acids of the invention are useful for  
 CC treating, diagnosing and monitoring progression of such diseases and  
 CC disorders as characterized by increased apoptotic cell-death of vascular  
 CC endothelial cells e.g. cardiac hypertrophy, myocardial infarction,  
 CC stroke, arteriosclerosis and heart failure. The present sequence  
 CC is one of the four Genbank sequences (A1761441.1) which are homologous to  
 CC the cDNA for human MIVR-1 and which are specifically disclaimed.  
 CC  
 XX  
 SQ Sequence 693 BP; 101 A; 205 C; 237 G; 149 T; 1 other;  
 Query Match 51.8%; Score 445.6; DB 24; Length 693;  
 Best Local Similarity 97.2%; Pred. No. 2,7e-82;  
 Matches 485; Conservative 0; Mismatches 10; Indels 4; Gaps 3;  
 QY 364 CGGCCCTTGGCCAGGGAGGCGCTTCCAGCGCTTCCAGCCACCTATCCGTACCTGCA 423  
 Db 693 CGGCCCTTGGCCAGGGAGGCGCTTCCAGCGCTTCCAGCCACCTATCCGTACCTGCA 634  
 QY 424 CACGAGATGCACTGGCCACCAACATCTGCTGTGAGAGGGGAGGCGCCACCTTAC 483  
 Db 633 CACGAGATGCA-CTGCGCGCCACCATCTGCTGTGAGAGGGGAGGCGCCACCTTAC 577  
 QY 484 CAGGCGCCCTGCAACCTTCCAGCTTGGGAGCCCGAGCACTGTAAGTGAACCGGGAG 543  
 Db 576 CAGGCGCCCTGCAACCTTCCAGCTTGGGAGCCCGAGCACTGTAAGTGAACCGGGAG 517  
 QY 544 TGGGTGCGGCGACCCCAAGAACCATCTGAGAGTGAATGATGATGTCAG 603  
 Db 516 TGGGTGCGGCGACCCCAAGAACCATCTTGAAGTGAATGATGTCAG 457  
 QY 604 CTGGGCGGCGCCCTGCGCCCGCCAGCACTGCGGATCAGCGCCACGTCATGCGGAGC 663  
 Db 456 CTGGGCGGCGCCCTGCGCCCGCCAGCACTGCGGATCAGCGCCACGTCATGCGGAGC 397  
 QY 664 GCGGCGGATGAGAGGGGCGCGCCACCTTACAGAGGATGATGCGGACATGACCGGGG 723  
 Db 396 GCGGCGGATGAGAGGGGCGCGCCACCTTACAGAGGATGATGCGGACATGACCGGGG 337  
 QY 724 TCCCTCTTCCAGACCAAGAGCACTGGGCGCGCCCTCTTCTGTA-GGGAGCCGGCT 782  
 Db 336 TCCCTCTTCCAGACCAAGAGCACTGGGCGCGCCCTCTTCTGTAAGGGAGCGGGT 277  
 QY 783 CCACGACACACATGCGCCCTTACAGAGCGGAGCGGATGAGGAGGAGGAGGATTA 842  
 Db 276 CCGCCACACACATGCGCCCTTACAGAGCGGAGCGGATGAGGAGGAGGAGGATTA 217  
 QY 843 ACAGAAAGGACACCCCTCTC 861  
 Db 216 ACAGAAAGGACACCCCTCTC 198

Search completed: March 17, 2003, 12:28:57  
 Job time : 238.386 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 12:18:54 ; Search time 41.4322 Seconds  
(without alignments)  
6373.035 Million cell updates/sec

Title: US-09-934-249-3  
Perfect score: 861  
Sequence: 1 atgacacgccttgatg95ggtc.....aacagaaagcaccctctc 861

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents-NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCtUS.COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	352.2	40.9	921	US-09-091-952A-7	Sequence 7, Appli
2	352.2	40.9	8065	US-09-091-952A-6	Sequence 6, Appli
3	284.2	33.0	867	US-09-091-952A-8	Sequence 8, Appli
4	47.8	5.6	8438	US-07-945-283-1	Sequence 1, Appli
5	46.2	5.4	15378	US-08-785-420-1	Sequence 1, Appli
6	45.4	5.3	424	US-08-476-176B-7	Sequence 7, Appli
7	45.4	5.3	424	US-08-127-721A-7	Sequence 7, Appli
8	45.4	5.3	424	US-08-485-246A-7	Sequence 7, Appli
9	44	5.1	1293	US-08-924-440-1	Sequence 1, Appli
10	44	5.1	44377	US-08-804-227C-7	Sequence 1, Appli
11	44	5.1	44377	US-08-804-198-1	Sequence 1, Appli
12	43.8	5.1	424	US-08-476-176B-5	Sequence 5, Appli
13	43.8	5.1	424	US-08-476-176B-9	Sequence 5, Appli
14	43.8	5.1	424	US-08-127-721A-5	Sequence 5, Appli
15	43.8	5.1	424	US-08-127-721A-9	Sequence 5, Appli
16	43.8	5.1	424	US-08-485-246A-5	Sequence 5, Appli
17	43.8	5.1	424	US-08-485-246A-9	Sequence 9, Appli
18	43.8	5.1	6834	US-09-194-905-7	Sequence 9, Appli
19	43.2	5.0	3705	US-08-474-379C-64	Sequence 64, Appli
20	43.2	5.0	3705	US-09-146-249A-64	Sequence 64, Appli
21	43.2	5.0	3705	US-08-206-188B-64	Sequence 64, Appli
22	43.2	5.0	152331	US-09-128-155-16	Sequence 16, Appli
23	42.8	5.0	4183	US-08-966-083-2	Sequence 2, Appli
24	42.8	5.0	4183	US-09-423-516-2	Sequence 2, Appli
25	42.2	4.9	1085	US-09-509-902A-3	Sequence 3, Appli
26	42.2	4.9	2793	US-08-209-747-1	Sequence 1, Appli
27	42.2	4.9	2793	US-08-458-298-1	Sequence 1, Appli

28	42.2	4.9	18596	4	US-09-312-468-11	Sequence 11, Appl	
c	29	41.6	4.8	30001	1	US-08-125-468-1	Sequence 1, Appl
c	30	41.6	4.8	30001	2	US-08-474-933-1	Sequence 1, Appl
c	31	41.4	4.8	693	4	US-09-402-515A-17	Sequence 17, Appl
32	41.2	4.8	377	2	US-08-332-766A-1	Sequence 1, Appl	
c	33	41	4.8	888	4	US-08-765-907A-2	Sequence 2, Appl
c	34	41	4.8	2277	1	US-08-676-967-5	Sequence 5, Appl
35	41	4.8	2277	1	US-08-676-974-5	Sequence 5, Appl	
36	41	4.8	2277	2	US-09-098-487-5	Sequence 5, Appl	
37	41	4.8	2888	4	US-08-765-907A-1	Sequence 1, Appl	
c	38	40.8	4.7	1743	3	US-08-665-259-20	Sequence 20, Appl
c	39	40.8	4.7	1743	3	US-08-762-500-78	Sequence 78, Appl
c	40	40.8	4.7	1974	3	US-08-762-500-19	Sequence 19, Appl
c	41	40.8	4.7	6803	3	US-08-665-259-1	Sequence 19, Appl
c	42	40.8	4.7	6803	3	US-08-762-500-19	Sequence 19, Appl
c	43	40.8	4.7	28804	2	US-08-592-874-1	Sequence 1, Appl
c	44	40.8	4.7	28804	3	US-09-096-942-2	Sequence 2, Appl
c	45	40.8	4.7	28804	3	US-09-096-867-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-091-952A-7  
Sequence 7, Application US/09091952A  
Patent No. 6A58532  
GENERAL INFORMATION:  
APPLICANT: Detera-Wadleigh, Sevilla D.  
Gershon, Elliot S.  
Badner, Judith A.  
Goldin, Lynn R.  
Berretini, Wade H.  
Yoshikawa, Takeo  
Sanders, Alan R.  
Estlerling, Lisa E.  
TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Maltic-Depressive Illness  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/091,952A  
FILING DATE: 19-Apr-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,278  
FILING DATE: 28-Oct-1996  
APPLICATION NUMBER: PCT/US97/19381  
FILING DATE: 28-Oct-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Timothy L.  
REGISTRATION NUMBER: 35,367  
REFERENCE/DOCKET NUMBER: 015280-297100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 921 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear  
 FEATURE: CDNA  
 NAME/KEY: CDS  
 LOCATION: 1...921  
 OTHER INFORMATION: Clone 22 coding region  
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
 US-09-091-952A-7

Query Match 40.9%; Score 352.2; DB 4; Length 921;  
 Best Local Similarity 73.1%; Pred. No. 1.4e-69;  
 Matches 482; Conservative 0; Mismatches 168; Indels 9; Gaps 2;

94 CAGAGCATGAGATGATGAGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGTATG 153  
 166 CCGGGCATCTTCAACTCGGAGCTGGAGTTCCGCAATCATCATCATCATCGTGGTGTATG 225  
 154 ATGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 213  
 226 ACGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 285  
 214 TTCATCAGCGCGGACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 273  
 286 TTCATCAGCGCGGACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 345  
 274 CTGTGCGCTTGGAGAGCAGAGTGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 333  
 346 CTGTGCGCTTGGAGAGCAGAGTGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 399  
 334 CCGCTGCGCGGACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 393  
 400 GCGCGGCTTGGAGAGCAGAGTGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 459  
 394 CCGCTGCGCGGACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 453  
 460 CCGCTGCGCGGACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 519  
 454 CTGTGCGCTTGGAGAGCAGAGTGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 513  
 520 CTGTGCGCGGAG 579  
 514 CCGGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 573  
 580 CCGGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 639  
 574 TTGCACTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630  
 640 TTGCACTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 699  
 631 AACTGCGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 690  
 700 AACTGCGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 759  
 691 AACTGCGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 749  
 760 AACTGCGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 818

RESULT 2  
 US-09-091-952A-6  
 Sequence 6 Application US/09091952A  
 Patent No. 6458532  
 GENERAL INFORMATION:  
 APPLICANT: Detera-Madleigh, Sevilla D.  
 Gerhson, Elliot S.  
 Badner, Judith A.  
 Goldin, Lynn R.  
 Berrettini, Made H.  
 Yoshikawa, Takeo  
 Sanders, Alan R.  
 Esterling, Lisa E.

TITLE OF INVENTION: Chromosomal Markers and Diagnostic  
 Tests for Manic-Depressive Illness

NUMBER OF SEQUENCES: 197  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94111-3834

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTED for Windows Version 2.0

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/091,952A  
 FILING DATE: 19-Apr-1999  
 CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/029,278  
 FILING DATE: 28-Oct-1996  
 APPLICATION NUMBER: PCT/US97/19381  
 FILING DATE: 28-Oct-1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Smith, Timothy L.  
 REGISTRATION NUMBER: 35,367  
 REFERENCE/DOCKET NUMBER: 015280-297100US

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8065 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: -  
 LOCATION: 1...8065  
 OTHER INFORMATION: Clone 22

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 116...1036  
 OTHER INFORMATION: Clone 22 coding region

FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 452...505  
 OTHER INFORMATION: alternatively spliced portion

FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 5595...5685  
 OTHER INFORMATION: amplified region for genotyping

US-09-091-952A-6  
 Query Match 40.9%; Score 352.2; DB 4; Length 8065;  
 Best Local Similarity 73.1%; Pred. No. 2.3e-69;  
 Matches 482; Conservative 0; Mismatches 168; Indels 9; Gaps 2;

94 CAGAGCATGAGATGATGAGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGTATG 153  
 281 CCGGGCATCTTCAACTCGGAGCTGGAGTTCCGCAATCATCATCATCATCGTGGTGTATG 340  
 154 ATGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 213  
 341 ACGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 400  
 214 TTCATCAGCGCGGACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 273  
 401 TTCATCAGCGCGGACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 460  
 274 CTGTGCGCTTGGAGAGCAGAGTGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 333



```

Db 461 CTGTGGCTTACAGACAGCGCCGACCGGGCTGG-----CGCCTGGAGATCATGCAT 514
Qy 334 CCGGCTGGGCGCCAGCAGCCGCTGGCGCGTGGCCCTTGGCCAGCGGAGCGCTTCCAC 393
Db 515 GCGCCGGGGTTCAGAGGATGATCAGACGCGCGTCTTCAATCCAGAGGATCGCTTCAGC 574
Qy 394 CGCTTCCAGCCGCTTATCCCTACCTGAGCAGAGATGACATGCCACCCACCATCTCG 453
Db 575 CGCTTCCAGCCGCTTATCCCTATGTGTGAGCAGAGATGATGATCTTCTCCACCATCTCC 634
Qy 454 CTGTCAAGCGGGAGAGCCGCCACCTTACGAGGCGCCCTGACCCCTCCAGCTTCGGGAC 513
Db 635 CTGTCCAGCGGTGAAGAGCCACCTCTTACCAAGGGGCGCTGACCCCTGACCTCCGGGAC 694
Qy 514 CCCGAGACGAGCTGGAACCTAACCGGAGTGTGCGCGCAGACCCCAAGAGAACATC 573
Db 695 CCTGACAGCAGATGAGAACTCAACCGAGATCGTGAAGGGCCCAACCGAACCATTA 754
Qy 574 TTGACAGTGAAGCTGATGATAGTGCAGG---CTGGGCGGCGCCCTGCGCCCGACGAGT 630
Db 755 TTTCACAGTGTATTAATAGATCTGTATGTATAGCGGGGCTCATGCGCCACCGAGCGC 814
Qy 631 AACTGGGCGATCAGCGCCACGCTGTACGCGCAGCGCGCGCATGAGAGGGCGCGCGCC 690
Db 815 AACTGGGCGATCAGTGAAGACACCTGACAGCAGTAAACGGGAGATGAGAGGGCGCACCGCC 874
Qy 691 ACCTACAGCGAGTGCATCGCCACTACCGGGGTCTCTTCACAGCAGCAGACGAG 749
Db 875 ACATACAGCGAGTGTATGGGCGCACCGCCGCTCTTCTTCATACACAGCGAG 933

```

## RESULT 3

```

US-09-091-952A-8
; Sequence 8, Application US/09091952A
; Patent No. 6458532

```

## GENERAL INFORMATION:

```

APPLICANT: Detera-Wadleigh, Sevilla D.
            Gershon, Elliott S.
            Badner, Judith A.
            Goldin, Lynn R.
            Berretini, Wade H.
            Yoshikawa, Takeo
            Sanders, Alan R.
            Esterling, Lisa E.
TITLE OF INVENTION: Chromosomal Markers and Diagnostic
                    Tests for Manic-Depressive Illness
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,952A
FILING DATE: 19-Apr-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029, 278
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: PCT/US97/19381
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 015280-297100US

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 867 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: cDNA
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..867
; OTHER INFORMATION: Clone 22 isoform 2 alternatively
; spliced coding region
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
;
US-09-091-952A-8

```

```

Query Match          33.0%; Score 284.2; DB 4; Length 867;
Best Local Similarity 68.7%; Pred. No. 1.6e-54;
Matches 453; Conservative 0; Mismatches 143; Indels 63; Gaps 2;

```

```

Qy 94 CAGAGCATGAGATCAGCAGACCTGAGTTGTTCAGATCATCATCATCTGTGTGATG 153
Db 166 CCGGGCATCTTCAACTGAGACTGGAGTGGCCCAATCATCATCATCTGTGTGATG 225
Qy 154 ATGATGATGTGTGTGTGTATACGTGCTGCTGAGACCACTAAGCTGTGACAGGTCC 213
Db 226 ACGGTGATGTGTGTGTGTATACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
Qy 214 TTTCATCAGCGCGGCGACAGCAGGCGGAGAGAGAGATGCGTCTCGAAGGATGC 273
Db 286 TTTCATCAGCGCGGCGACAGCAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 329
Qy 274 CTGTGGCCCTTGAGAGACAGTGTCAAGGCAAGCAATCCAGAGCGGAGTCTAGGCC 333
Db 330 -----GCGGCAATCATGCAT 345
Qy 334 CCGGCTGGGCGCCAGCAGCGCTGCGCGTGGCGCCCTTCCAGCGGAGCGCTTCCAC 393
Db 346 GCGCGGCGGTCCAGGAGAGGTTTCAAGCGCGCGTCTTCAATCCAGAGGATCGCTTACG 405
Qy 394 CGCTTCCAGCGCCACTTATCCCTATGTGTGACAGCAGATGATGATCTTCTCCACATCTCC 465
Db 406 CGCTTCCAGCGCCACTTATCCCTATGTGTGACAGCAGATGATGATCTTCTCCACATCTCC 465
Qy 454 CTGTCAAGCGGGAGAGACCCCACTTACGAGGCGCCCTGACCCCTCCAGCTTCGGGAC 513
Db 466 CTGTCCAGCGGTGAAGAGCCACCTCTTACCAAGGGGCGCTGACCCCTGACCTCCGGGAC 525
Qy 514 CCCGAGACGAGCTGGAACCTAACCGGAGTGTGCGCGCAGACCCCAAGAGAACATC 573
Db 526 CCTGACAGCAGATGAGAACTCAACCGAGATCGTGAAGGGCCCAACCGAACCATTA 585
Qy 574 TTGACAGTGAAGCTGATGATAGTGCAGG---CTGGGCGGCGCCCTGCGCCCGACGAGT 630
Db 586 TTGACAGTGAATTAATAGATCTGTATGTATAGCGGGGCTCATGCGCCACCGAGCGC 645
Qy 631 AACTGGGCGATCAGCGCCACGCTGTACGCGCAGCGCGCGCATGAGAGGGCGCGCGCC 690
Db 646 AACTGGGCGATCAGTGAAGACACCTGACAGCAGTAAACGGGAGATGAGAGGGCGCACCGCC 705
Qy 691 ACCTACAGCGAGTGCATCGCCACTACCGGGGTCTCTTCACAGCAGCAGACGAG 749
Db 706 ACATACAGCGAGTGTATGGGCGCACCGCCGCTCTTCTTCATACACAGCGAG 764

```

## RESULT 4

```

US-07-945-283-1/c
; Sequence 1, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:

```

```

APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
TITLE OF INVENTION: Involving The Ep0 and LfT Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS:
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4011 ext.513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8438 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudorabies virus
FEATURE:
NAME/KEY: CDS
LOCATION: 622..6495
FEATURE:
NAME/KEY: variation
LOCATION: replace(1099, "g")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1267, "t")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1381, "c")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1566, "c")
FEATURE:
NAME/KEY: variation
LOCATION: replace(7010, "g")
US-07-945-283-1

Query Match
Best Local Similarity 5.6%; Score 47.8; DB 1; Length 8438;
Matches 189; Conservative 0; Mismatches 217; Indels 1; Gaps 1;
QY 353 GCCTGGCCGTCGCGCCCTTCCGCGAGCGGAGCGCTTCACCGCTTCACCGCCACCGTATC 412
DB 6181 GCCCGGGCGGGCGCGGCTTCGAGACCCCGCGCCACACGACGACGACGCGGTGC 6122
QY 413 CGACTGACGACGAGATGACCTGACCGACCGACCATCTGCTGTACAGAGGGAGAGC 472
DB 6121 CGCAGCGCGGACGCGGCGGGGCGCGCTTCC -CAGCTCCCGCGCGCGCGCGGTCT 6063
QY 473 CCCACCTTACAGGGGCGCGCTTCACCTCAAGCTTCGGGAGCCCGAGCGAGCGTGAAC 532
DB 6062 CGGCGCGCGGGCGCGCGCGCGCTTCGCGCGAGGAGCGCGCGGTCTCTACCATCAC 6003

```

```

QY 533 TGACCGGAGTGTGCTGCGCGACCCCAACAGAACATCTTGACAGTACCTGATG 592
DB 6002 AGACCGGAGCGCGCGCTTCGACGAGCCCGCGCGGAGGCGGCTGATCTCGGCCGC 5943
QY 593 ATAGTCCAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5943
DB 5942 TTGGCGCGCGGAGCGCGCGCTCTCTCAACCCCAACTCCAGCTCCAGCTCCAC 5883
QY 653 GCTACGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 712
DB 5882 TCGCGGTGAGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5823
QY 713 ACTACCGGCGCGCTCTCTCTTCACAGCAGCAGCAGCAGTGGCGCGCG 759
DB 5822 GGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5776

```

```

RESULT 5
US-08-785-420-1/c
Sequence 1, Application US/08785420
Patent No. 6001976
GENERAL INFORMATION:
APPLICANT: MacLennan, David H
APPLICANT: O'Brien, Peter J.
TITLE OF INVENTION: DIAGNOSIS FOR PORCINE MALIGNANT
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Bell, Seltzer, Park & Gibson
STREET: P.O. Drawer 34009
CITY: Charlotte,
STATE: No. 6001976th Carolina 28234
COUNTRY: U.S.A.
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,420
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/249,388
FILING DATE:
APPLICATION NUMBER: 15-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Layton, Jr., Samuel G.
REGISTRATION NUMBER: 22807
REFERENCE/DOCKET NUMBER: 3477-73
TELECOMMUNICATION INFORMATION:
TELEPHONE: 704-377-1561
TELEFAX: 704-377-1561
TELEX: 57-5102
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 15378 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: Porcine RYR1 Gene
POSITION IN GENOME:
UNITS: bp
US-08-785-420-1

Query Match
5.4%; Score 46.2; DB 3; Length 15378;

```



Query Match	5.38;	Score 45.4;	DB 3;	Length 424;
Best Local Similarity	47.48;	Pred. NO. 0.11;		
Matches 136;	Conservative	0.0;	MicroMatch	0.0

[illegible]

RESULT 8  
 US-08-485-246A-7  
 Sequence 7, Application US/08485246A  
 Patent No. 6072035  
 GENERAL INFORMATION:  
 APPLICANT: Hardman, No. 6072035man  
 APPLICANT: Kolbinger, Frank  
 APPLICANT: Saldanha, Jose  
 TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
 TITLE OF INVENTION: Immunoglobulin isotype  
 NUMBER OF SEQUENCES: 55  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: No. 6072035artis Patent Department  
 STREET: 59 Route 10  
 CITY: East Hanover  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07936-1080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/485,246A  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/127,721  
 FILING DATE: 27-SEPTEMBER-1993  
 APPLICATION NUMBER: US 07/952,802  
 FILING DATE: 25-SEPTEMBER-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: NO. 6072035ak, Henry P.  
 REGISTRATION NUMBER: 33,200  
 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (908) 277-5110  
 TELEFAX: (908) 277-4306

```

? INFORMATION FOR SEQ ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 424 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 22..402
? FEATURE:
? NAME/KEY: mat_peptide
? LOCATION: 82..402
? OTHER INFORMATION: /product="light chain variable
? region C21-I2"
US-08-485-246A-7

```

Query Match	5.38;	Score 45.4;	DB 3;	Length 424;
Best Local Similarity	47.48;	Pred. No. 0.11;		
Matches 136;	Conservative	0;	Mismatches 171	

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

RESULT 9  
 US-08-924-440-1  
 ; Sequence 1, Application US/08924440  
 ; Patent No. 5871550  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Erits et al.  
 ; TITLE OF INVENTION: MUTANT THERMONOSPORA SPP. CELLULOSE  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genencor International, Inc.  
 ; STREET: 925 Page Mill Road  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ;  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Fastseq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/924,440  
 ; FILING DATE: August 27, 1997  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ;  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Christopher L. Stone  
 ; REGISTRATION NUMBER: 35,656  
 ; REFERENCE/DOCKET NUMBER: GC368  
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 846-7555  
 TELEFAX: (650) 845-6504  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1293 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-924-440-1

Query Match 5.1%; Score 44; DB 2; Length 1293;  
 Best Local Similarity 47.5%; Pred. No. 0.3;  
 Matches 131; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 551 GCGCACCCCAAGCAAGCACTCTTCGACAGTGCATGATGTGCCAGGCTGGGCG 610  
 DB 263 GCGAACCCCAACCACTGCACCAATCAACAGGCGCCCTGCGACCAAGCTCCGAGCGGGGCG 322  
 QY 611 GCGCCCTGCGCCCGCCAGTACCTGCGGCGATCAGCGCCAGCTACGCGAGCGGCGGCG 670  
 DB 333 GCGCCGCGGCTCCCGAAGTCCAGTCTGCGGATCCAGCTTCGCGAGCAAGCGCA 382  
 QY 671 GCATGAGAGGGCGCGCCCGCCACCTACAGAGTATCGGCCACTACCGGGGCTCTCTCT 730  
 DB 383 CGGTGAGGCGGTACGCAAGTCCAGTCTGCGGATCCAGCTTCGCGAGCAAGCGCA 442  
 QY 731 TCCAGCACCAAGCAAGCACTGAGGCGGCTCTCTTCGCTGAGAGGAGCCGCGCTCCACACA 790  
 DB 443 ACCCGCTCCAACTGCGCGGATAGACACCCAGCGGATCCAGTGGTTCGACCACTGCTGA 502  
 QY 791 CACACATCGCGCCCTAGAGAGCGAGCCATCTGGA 826  
 DB 503 CCGACAGCTCGCTGAGACGCGCTGCGCTACGACTGGA 538

RESULT 10  
 US-08-804-227C-7  
 Sequence 7, Application US/08804227C  
 Patent No. 5876991

GENERAL INFORMATION:  
 APPLICANT: DeHoff, Bradley S.  
 APPLICANT: Kustross, Stuart A.  
 APPLICANT: Rostock, Paul R., Jr.  
 APPLICANT: Sutton, Kimberly L.  
 TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: THOMAS G. PLANT 1501  
 STREET: LILLY CORPORATE CENTER  
 CITY: INDIANAPOLIS  
 STATE: IN

COUNTRY: USA  
 ZIP: 46285  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: ASCII(DOS) Text only  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/804.227C  
 FILING DATE: February 21, 1997

CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plant, Thomas, G.  
 REGISTRATION NUMBER: 35,784  
 REFERENCE/DOCKET NUMBER: X-8231  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 317-276-2459  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4437 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single

TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 NAME/KEY: CDS  
 LOCATION: 350..14002  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 14046..20036  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 20110..31284  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 31329..36071  
 NAME/KEY: CDS  
 LOCATION: 36155..41830  
 US-08-804-227C-7

Query Match 5.1%; Score 44; DB 2; Length 4437;  
 Best Local Similarity 43.5%; Pred. No. 0.65;  
 Matches 200; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 330 GCGCCGCGCTCGGCGCCACGACCGCGCTGCGCGCTTCGCGCCAGCGGAGCGCTT 389  
 DB 40615 GCGCGCGCGCGCGCGCGCGCGCGAGACACCCCGCGCGCGCTTCGCGCGCGCGCG 40674  
 QY 390 CCACGCGCTTCAGCCCACTTATCCGCTGACGACGAGATGACCTGCGCACCACT 449  
 DB 40675 CGAGAGACCGCGCGCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 40734  
 QY 450 CTGCGTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 509  
 DB 40735 CCAGTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 40794  
 QY 510 GGACCCCGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 569  
 DB 40795 GCTGCGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 40854  
 QY 570 CATCTTCGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 629  
 DB 40855 CACCTTCGCTTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 40914  
 QY 630 TAACTGCGGAG 689  
 DB 40915 CGAGCGGAG 40974  
 QY 690 CACCTACAG 749  
 DB 40975 CTCCTTCG 41034  
 QY 750 TGGGCGCGCGCTTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 789  
 DB 41035 CCTGCGCGCGCTTCG 41074

RESULT 11  
 US-08-804-198-1  
 Sequence 1, Application US/08804198  
 Patent No. 5945320  
 GENERAL INFORMATION:  
 APPLICANT: Burgett, Stanley G.  
 APPLICANT: Kustross, Stuart A.  
 APPLICANT: Rao, Nagaraja R.  
 APPLICANT: Richardson, Mark A.  
 APPLICANT: Rostock, Paul R., Jr.  
 TITLE OF INVENTION: PLATENOLOLIDE SYNTHASE GENE  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PAUL R. CANTRELL 1138  
 STREET: LILLY CORPORATE CENTER  
 CITY: INDIANAPOLIS  
 STATE: IN

```

1 COUNTRY: USA
2 ZIP: 46285
3
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: Floppy disk
6 COMPUTER: Macintosh
7 OPERATING SYSTEM: Macintosh 7.0
8 SOFTWARE: Microsoft Word 5.1
9
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER: US/08/804,198
12 FILING DATE:
13 CLASSIFICATION: 435
14 ATTORNEY/AGENT INFORMATION:
15 NAME: CANTRELL, PAUL R.
16 REGISTRATION NUMBER: 36,470
17 REFERENCE/DOCKET NUMBER: P9113
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE: 317-276-3885
20
21 INFORMATION FOR SEQ ID NO: 1:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 44377 base pairs
24 TYPE: nucleic acid
25 STRANDEDNESS: single
26 TOPOLOGY: linear
27 MOLECULE TYPE: DNA (genomic)
28
29 FEATURE:
30 NAME/KEY: CDS
31 LOCATION: 350..14002
32 FEATURE:
33 NAME/KEY: CDS
34 LOCATION: 14046..20036
35 FEATURE:
36 NAME/KEY: CDS
37 LOCATION: 20110..31284
38 FEATURE:
39 NAME/KEY: CDS
40 LOCATION: 31329..36071
41 FEATURE:
42 NAME/KEY: CDS
43 LOCATION: 36155..41830
44
45 US-08-804-198-1

```

Query Match	Similarity	5.1%	Score	44:	DB	2:	Length	44377:
Best Local	Similarity	43.5%	Pred.	No.	0.65:			
Matches	200:	Conservative	0:	Mismatches	260:	Indels	0:	Gaps
QY	330	CGCCCCGCTGGCCACCAGACCGCTGGCCGCTGGCCGCTTTCGCCCAACGGAGCGCTT	389					
Db	40615	CGCCCCGCGCGCGCGCGCGAGAACCCGCCACCGCCCGCGCGCTTCCTGGCGGGCGCC	40674					
QY	390	CCACCGCTTCCAGCCGCCACCTATCCGTACTCTGAGCAGCAGATCGACCTGCCACCACT	449					
Db	40675	CGAGACCGCGCGCGCGCTGGCCCTTGAACTGGTGTCTCGCCACGTGGCGGGCTCTCGG	40734					
QY	450	CTCGCTGTTCAGACGGGGAGAGACCCGCCAACCCTACCAAGGGGCCCTCACCCTTCAGCTTG	509					
Db	40735	CCACTCCGAGAGACGCCGGGTGAGACGCCCGGCCCGCCCTTCGGGAACCTGGGTTGACATC	40794					
QY	510	GGACCCCGAGCAGCAGCTGTGAACCTAACCCGGAGTGTGGTGGCGGCACCCCAACAGAAC	569					
Db	40795	GCTCCGCGCGGTGGCGCTGGCGCGCGCGCTGGCCAGAGCACCGGGCTTGAACTGGCCGG	40854					
QY	570	CATCTTCACACAGTAGCACTGATGATATGTATGTCCAGGCTGGGGCGGCCCTGCCCCAGCAG	629					
Db	40855	CACCGTCGTTCTTCAGACACAGAGAACCCCAACCGCGCTGGGCCACACATCGTGGCGCTCGC	40914					
QY	630	TAACTGGGGCATCAGCGCCACGCTGCTACGGGCAAGCGGGCGCATAGAGGGCGCGCGCC	689					
Db	40915	CGACCGGGGGAGCCCGCGGCCCCCAAGAGGGGCAACGCGCTGGGGCGAGCGGCTGTTCCG	40974					
-QY	690	CACCTACAGAGGAGTCATCTGGCCACTACCGGGGGTCTCTTCACAGACACAGAGCAG	749					
Db	40975	CTCTCTCCGCGCGCGCGCTGCAACACGCGCAGGTGCAACGAGTCTGTGTGACATGATGCCGA	41034					

Qy 750 TGGCCCGCCCTCCCTTGTGGAGGGGACCGGCTCCACCAC 789  
| | | | | | | | | | | | | | | | | | | | | |  
Db 41035 CCTGGCGGGCTTCGGGCCGCTACTCCCGGAGCACCAC 41074

```

RESULT 12
US-08-476-176B-5
Sequence 5, Application US/08476176B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
NUMBER OF INVENTION: Immunoglobulin isotype
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..402
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 82..402
OTHER INFORMATION: /product= "light chain variable
OTHER INFORMATION: region C21-L1"
US-08-476-176B-5

```

Query Match	5.13;	Score 43.8;	DB 2;	Length 424;
Best Local Similarity	47.08;	Pred. No. 0.26;	Mismatches 152;	Indels 0;
Matches 135;	Conservative	0;	Gaps	0;
QY 483	CCAGGCGCCCTGCACCCCTTCGAGTTCGGGACCCCGAGCAGCAGGCTGGAATCTGAACCCGGGA	542		
Db 96	CCAGAGCCCTGGACACCTGAGCCTGAGCCCGGGGAGAGGGGCCACCTGACCTGCAGGGCC	155		
QY 543	GTCGGTGGCGCCACACCCCAACAGAACCATCTTCGACAGTGAACGTGATGATAGTCAG	602		
Db 156	CAGCCAGAGCATTCGGACACCAACATCTGTTACGACGAGAAAGCCCGCCACGGCCCCAG	215		

Matches 135; Conservative 0; Mismatch 0; Local similarity 47.08; FICU: 0

```

; LOCATION: 82.1402
; OTHER INFORMATION: /product= "light chain variable

```

Query Match	5.18;	Score 43.8;	DB 3;	Length 424;
Best Local Similarity	47.0%;	Pred. No. 0.26;		
Matches 135; Conservative	0;	Mismatches 152;	Indels 0;	Gaps 0;

```

;
; FEATURE:
; NAME /REV.
; CDS

```







```
Sequence 55, Application US/09796753
Publication No. US20030027998A1
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-227-999
CURRENT APPLICATION NUMBER: US/09/796,753
PRIORITY FILING DATE: 2001-03-01
PRIORITY APPLICATION NUMBER: 09/183,175
PRIORITY FILING DATE: 1998-10-30
PRIORITY APPLICATION NUMBER: 09/223,094
PRIORITY FILING DATE: 1998-12-30
PRIORITY APPLICATION NUMBER: 09/223,546
PRIORITY FILING DATE: 1998-12-30
PRIORITY APPLICATION NUMBER: 09/224,246
PRIORITY FILING DATE: 1998-12-30
PRIORITY APPLICATION NUMBER: 09/259,388
PRIORITY FILING DATE: 1999-02-26
PRIORITY APPLICATION NUMBER: 60/122,458
PRIORITY FILING DATE: 1999-03-01
PRIORITY APPLICATION NUMBER: 09/312,359
PRIORITY FILING DATE: 1999-05-14
PRIORITY APPLICATION NUMBER: 09/336,536
PRIORITY FILING DATE: 1999-06-18
PRIORITY APPLICATION NUMBER: 09/342,687
PRIORITY FILING DATE: 1999-06-29
PRIORITY APPLICATION NUMBER: 09/345,464
PRIORITY FILING DATE: 1999-06-30
PRIORITY APPLICATION NUMBER: 09/365,164
PRIORITY FILING DATE: 1999-07-30
PRIORITY APPLICATION NUMBER: 09/399,723
PRIORITY FILING DATE: 1999-09-20
PRIORITY APPLICATION NUMBER: 09/409,634
PRIORITY FILING DATE: 1999-09-30
PRIORITY APPLICATION NUMBER: 09/471,179
PRIORITY FILING DATE: 1999-12-23
PRIORITY APPLICATION NUMBER: 09/474,071
PRIORITY FILING DATE: 1999-12-29
PRIORITY APPLICATION NUMBER: 09/474,072
PRIORITY FILING DATE: 1999-12-29
PRIORITY APPLICATION NUMBER: 09/514,010
PRIORITY FILING DATE: 2000-02-25
PRIORITY APPLICATION NUMBER: 09/516,745
PRIORITY FILING DATE: 2000-03-01
PRIORITY APPLICATION NUMBER: 09/572,002
PRIORITY FILING DATE: 2000-05-14
PRIORITY APPLICATION NUMBER: 09/597,993
PRIORITY FILING DATE: 2000-06-19
PRIORITY APPLICATION NUMBER: 09/599,596
PRIORITY FILING DATE: 2000-06-22
PRIORITY APPLICATION NUMBER: 09/630,334
PRIORITY FILING DATE: 2000-07-31
PRIORITY APPLICATION NUMBER: 09/606,565
PRIORITY FILING DATE: 2000-06-29
PRIORITY APPLICATION NUMBER: 09/606,317
PRIORITY FILING DATE: 2000-06-29
PRIORITY APPLICATION NUMBER: 09/665,666
PRIORITY FILING DATE: 2000-09-30
PRIORITY APPLICATION NUMBER: 09/677,751
PRIORITY FILING DATE: 2000-09-30
NUMBER OF SEQ ID NOS: 162
SEQ ID NO 55
LENGTH: 969
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (6)...(761)
US-09-796-753-55
```

```
Query Match      87.7%: Score 755.2; DB 9; Length 969;
Best Local Similarity 99.6%: Pred. No. 7.7e-162;
Matches 757; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 102 GGAGATCAAGGAGCTGGAGTTTGTTCACATCATCATCGTGGTGGATGGTGCAT 161
Db 2 GGAGATGGCGGGAGCTGGAGTTTGTTCACATCATCATCGTGGTGGATGGTGCAT 61
QY 162 GGTGGTGGTGGATCAGTGCCTCTGTGAGCCACACACAGGTGTGTGACGGTCCCTTCATG 221
Db 62 GGTGGTGGTGGATCAGTGCCTCTGTGAGCCACACACAGGTGTGTGACGGTCCCTTCATG 121
QY 222 CCGGCACAGCCAGGGGGGAGAGAGAAAGATCCCTGTCTTCAGAGATGCTGTGACC 281
Db 122 CCGGCACAGCCAGGGGGGAGAGAGAAAGATCCCTGTCTTCAGAGATGCTGTGACC 181
QY 282 CTGGAGAGGACAGTGTAGAGCAAGCAAGCAATCCCAAGCCGAGTCTACGCCGCCCTCG 341
Db 182 CTGGAGAGGACAGTGTAGAGCAAGCAAGCAATCCCAAGCCGAGTCTACGCCGCCCTCG 241
QY 342 GCCCACCAGCCGCTGGGCCCTGCCCTTCGCCAGGGGAGAGCCCTTCACACCGTTCCA 401
Db 242 GCCCACCAGCCGCTGGGCCCTGCCCTTCGCCAGGGGAGAGCCCTTCACACCGTTCCA 301
QY 402 GCCCAGCTATCCGTACTCTGACGACAGAGATGACCTGCCACCCACATCTGCTTCAGA 461
Db 302 GCCCAGCTATCCGTACTCTGACGACAGAGATGACCTGCCACCCACATCTGCTTCAGA 361
QY 462 CGGGAGAGGAGCCGCCACCCCTACAGGGGCCCTTCGCCAGCCCTTCGGGAGCCCGAGCA 521
Db 362 CGGGAGAGGAGCCGCCACCCCTACAGGGGCCCTTCGCCAGCCCTTCGGGAGCCCGAGCA 421
QY 522 GCAGCTGGAACCTGAACCGGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 581
Db 422 GCAGCTGGAACCTGAACCGGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 481
QY 582 TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 641
Db 482 TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
QY 642 CAGCGCCACGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 701
Db 542 CAGCGCCACGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601
QY 702 GGTGATCGGCGCTACCGCGGGGTCTCTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGC 761
Db 602 GGTGATCGGCGCTACCGCGGGGTCTCTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGC 661
QY 762 CTTCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 821
Db 662 CTTCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 721
QY 822 CTGGAGCAAGAGAGAGATTAACAGAAAGAGACACCTCTTC 861
Db 722 CTGGAGCAAGAGAGAGATTAACAGAAAGAGACACCTCTTC 761

RESULT 4
US-10-098-841-71
Sequence 71, Application US/10099841
Publication No. US20020197679A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunqing
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
```

Query Match	Best Local Similarity	99.2%	Score 742.4	DB 9	length 1066
Matches 746	Conservative	0	Mismatches 6	Indels 0	Gaps 0
QY 110	CGAGGCTGGAGTGTGTTGATCATCATCATCATGTCGTGTGATATATATGATGATGATG	169			
Db 113	CGGACGACGATCTTGTTCAGATCATCATCATCATCTGTGTGTATGATGATGATGATGATG	172			
QY 170	TGATCACTGTCCTGGCTGAGACCCACTCAACGCTGTGTGACAGGATCTTCATAGCGGGACA	229			
Db 173	TGATCACTGTCCTGGCTGAGACCCACTCAACGCTGTGTGACAGGATCTTCATAGCGGGACA	232			
QY 230	GCAGAGGGGGGAGAGAGAAATGTCCTGTCTCAGAAAGATCTCTGTGGCCCTGGAGA	289			
Db 233	GCAGAGGGGGGAGAGAGAAATGTCCTGTCTCAGAAAGATCTCTGTGGCCCTGGAGA	292			
QY 290	GCACAGTGTCAAGGCACACGAATCCCAAGGCGAGGTCTTACGCCCGCTCGGCCACCG	349			
Db 293	GCACAGTGTCAAGGCACACGAATCCCAAGGCGAGGTCTTACGCCCGCTCGGCCACCG	352			
QY 350	ACCGGCTTGGCGGTGGCGCCCTTGCCCAAGGGAGGCGCTTCCACGCTTCCAGGCCACT	409			
Db 353	ACCGGCTTGGCGGTGGCGCCCTTGCCCAAGGGAGGCGCTTCCACGCTTCCAGGCCACT	412			
QY 410	ATCCGACTCTGACACAGAGATGACATGCACACCCACATCTCGTGTGACAGCGGGAGG	469			
Db 413	ATCCGACTCTGACACAGAGATGACATGCACACCCACATCTCGTGTGACAGCGGGAGG	472			
QY 470	AGCCCCACCTCTCAAGAGGCCCTCTGACCCCTCAAGCTTGGGAGCCCGACAGCAGCTGG	529			
Db 473	AGCCCCACCTCTCAAGAGGCCCTCTGACCCCTCAAGCTTGGGAGCCCGACAGCAGCTGG	532			
QY 530	AACGTGAACCGGAGTGTGTTGGCGACCCCAACACAAACATCTTCAGAGTACCTGA	589			
Db 533	AACGTGAACCGGAGTGTGTTGGCGACCCCAACACAAACATCTTCAGAGTACCTGA	592			
QY 590	TGGATAGTCCAGAGCTGGGGCGCCCTGCCCCCAAGCACTTAATCTGGGATCAGCGCA	649			
Db 593	TGGATAGTCCAGAGCTGGGGCGCCCTGCCCCCAAGCACTTAATCTGGGATCAGCGCA	652			
QY 650	CGTGTACGAGGAGCGGGGCGCATGAGAGGGCGCGGCCCACTACAGCGAGGTCACTG	709			
Db 653	CGTGTACGAGGAGCGGGGCGCATGAGAGGGCGCGGCCCACTACAGCGAGGTCACTG	712			
QY 710	GCACACTACCGGGGTCTCTCTTCACAGACAGAGAGAGTGTGGCGCCCTCTTGTCTGG	769			
Db 713	GCACACTACCGGGGTCTCTCTTCACAGACAGAGAGAGTGTGGCGCCCTCTTGTCTGG	772			
QY 770	AGGGAGCCCGGCTGCACACACACACATCGCGCCCTAGAGAGGCGACCATCTGGAGCA	829			

Db	773	AGGGGACCCGGCTCCACACACACATCGGCGCCCTAGAGAGCGGACGCATCTTGAGCA	832
Oy	830	AAGGAGAGATTAACAGAAAGAGACACCCCTGC	861
Db	833	AAGGAGAGATTAACAGAAAGAGACACCCCTCTC	864

RESULT 5  
US-09-934-249-12  
; Sequence 12, Application US/09934249  
; Patent No. US20020115081A1  
; GENERAL INFORMATION

APPLICANT: Lee, Richard T.  
APPLICANT: Landschulz, Katherine T.  
APPLICANT: Turi, Thomas G.  
APPLICANT:

APPLICANT: Kennedy, Scott P.  
TITLE OF INVENTION: DISCRETE

TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS  
FILE REFERENCE: P0738/7001/ERP/KA

CURRENT FILING DATE: 2001-08-21  
 PRIOR APPLICATION NUMBER: US 60/032,471

PRIOR FILLING DATE: 2000-08-22  
 NUMBER OF SEQ ID NOS: 17  
 SOFTWARE: FASTA

```

; SEQ ID NO 12
; LENGTH: 878
; windows version 3.0

```

ORGANISM: Mus Musculus  
FEATURE:

```

; NAME/KEY: CDS
; LOCATION: (20)...(841)
US-09-934-249-12

```

Query Match	71.1%;	Score 612.4;	DB 10;	length 878;
Best local Similarly	86.5%;	Pred. No. 1.4e-129;		
Matches 701; Conservative	0;	Mismatch		

	CACCGCTTCATGGG	GCTCAACA	GCACCCGCCGC	CCCCC	CCCCCCC	CCCCCGAGAGCCAAATTCTTTTTC	67
4							

II CATTGGCTTGATGGGGGTCACAGGCGACCGCGCCGCCGCCGCCGGCGGCGACAGCCCATGTCCTCC 70

[illegible]

124 GTTCAGATCATCATCGTGGTGCTGATGATGATGATGTTGGTGCCTCATTGACCTGCT  
130

131 GTCAATCGTGTGTCATCGTGTGCTAGTCATGATGATGCTGCTATGATTACGTGCTG 190

[illegible]

244 AGAGAAGATGCCCTGTCTCAGAGGATGCCmctccggcgcgccc-  
tcccccccccctccatcagccgcacacgccaggccagagc 250

251 AGAGACGATGGAAGATGCTCTGGCCCTCAGAGATACGGTGTACCT - 300

304 AACGGAATCCGACGAGCCGCTACGCCCCGCCCTCGGCCACCGGACCGGCTGCGCGTG 363

364 CCGCCCTTCGACCCAGCGGCATCCCCC  
-CCCAACCAGGCCACAGGTCTATGCCCGCCCTCGGGCCACTGACCGACTCGCTGTG 367

368 CCCCCCTTCATCCAGCG-----AGCGGATTCCAACCCACCCTTGGCAG 423

424 CACGAGATGACCTGCCACCCACCATCTGCTGCAGACGGGGGAGAGCCGCCACCTAC 483

415 CACGAAATTCGCCACCCACCATCTCACTGTCGTGATGGGAGAGAGCCGCCACCCCTAC 478

Db 479 CAGGGCCCTCAGCCCTCAGCTACGGGACCTGAGCAACAGCTGAGCTGAACCGGAA 538  
Qy 544 TCGGTGGCGGACCCCAAGAAACATCTTCGACACTGACATGATGATGATGATGATG 603  
Db 539 TCTGTGGCGGACCCCTTAACCGGACATCTTCGACATGATGATGATGATGATGATG 598  
Qy 604 CTGGGCGGCGCCTGCGCCCGCAGCAGTAACTCGGGCATCAGCGCCAGCTGCTACGGCAG 663  
Db 599 CTGGGCGGCGCCTGCGCCCGCAGCAGTAACTCGGGCATCAGCGCCAGCTGCTACGGCAG 658  
Qy 664 GCGGCGGCGCATGAGAGGGGCGCGCCGACCTACAGCGAGGTCATCGCCACTACCGGGAG 723  
Db 659 GGTGGGCGCATGAGAGGGGCGCGCCGACCTACAGCGAGGTCATGCGCCACTACCGCTGC 718  
Qy 724 TCGTCCTTCAGCAGCAGCAGAGAGAGTGGGCGCCCTCTGCTGAGAGGAGCGCGCTC 783  
Db 719 TCGTCCTTCAGCAGCAGCAGAAAGTAAAGGCGCCATCTCTGCTAAGAGGGAGCCGCGCTC 778  
Qy 784 CACCAACACACATCGCGCCCTAGAGAGC 813  
Db 779 CATCACTGCACTTGCCCTCAGTGAAGAC 808

RESULT 6  
US-10-000-256A-32  
; Sequence 32, Application US/10000256A  
; Publication No. US20030039983A1  
; GENERAL INFORMATION:  
; APPLICANT: Sun, Yonming  
; APPLICANT: Recipon, Herve  
; APPLICANT: Chen, Sei-Yu  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
; FILE REFERENCE: DEX-0259  
; CURRENT APPLICATION NUMBER: US/10/000.256A  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 60/244,782  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 240  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32  
; LENGTH: 1583  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-000-256A-32

Query Match 63.2%; Score 544.4; DB 9; Length 1583;  
Best Local Similarity 99.8%; Pred. No. 3.6e-114;  
Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 316 GAGCGCGAGTCTACGCGCCGCTGCGCCACGACCGCGTGGCGCCCTTCGCGC 375  
Db 825 GCGCGCGAGTCTACGCGCCGCTGCGCCACGACCGCGTGGCGCCCTTCGCGC 884  
Qy 376 CAGCGGAGCGGCTTCACCGGCTTCAGCGCCACTTCGCTGCTGCTGCTGCTGCTGCTG 435  
Db 885 CAGCGGAGCGGCTTCACCGGCTTCAGCGCCACTTCGCTGCTGCTGCTGCTGCTGCTG 944  
Qy 436 CTGCGACCCACCATCTCGCTGTACAGCGGGGAGAGCCCGCCACCTTACAGGGCCCTGCG 495  
Db 945 CTGCGACCCACCATCTCGCTGTACAGCGGGGAGAGCCCGCCACCTTACAGGGCCCTGCG 1004  
Qy 496 ACCCTCAGCTTCGGGACCCGAGCAGCAGCTGAAGTGAACCGGAGTGGTGGCGCA 555  
Db 1005 ACCCTCAGCTTCGGGACCCGAGCAGCAGCTGAAGTGAACCGGAGTGGTGGCGCA 1064  
Qy 556 CCCCCAAGAGACCATCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 615  
Db 1065 CCCCCAAGAGACCATCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1124  
Qy 616 TGCCCCCAGCAGTAACTCGGGCATCAGCGCCAGCTGCTACGCGAGGCGGCGCATG 675  
Db 1124 TGCCCCCAGCAGTAACTCGGGCATCAGCGCCAGCTGCTACGCGAGCGGCGGCATG 675

Db 1125 TGCCCCCAGCAGTAACTCGGGCATCAGCGCCAGCTGCTACGCGAGCGGCGGCATG 1184  
Qy 676 GAGGGCGCGCGCCACCTACAGGAGTCAATCGGCACTACCGGGGCTCTCTTCAG 735  
Db 1185 GAGGGCGCGCGCCACCTACAGGAGTCAATCGGCACTACCGGGGCTCTCTTCAG 1244  
Qy 736 CACGAGCAGCAGTGGGCGCGCCCTCTCTGCTGAGAGGAGCGCGCTCAGCAGCAGC 795  
Db 1245 CACGAGCAGCAGTGGGCGCGCCCTCTCTGCTGAGAGGAGCGCGCTCAGCAGCAGC 1304  
Qy 796 ATCGCGCCCTTAGAGAGCGGACCATCTGAGAGCAAGAGATTAACAGAAAGAGAC 855  
Db 1305 ATCGCGCCCTTAGAGAGCGGACCATCTGAGAGCAAGAGATTAACAGAAAGAGAC 1364  
Qy 856 CCGTC 861  
Db 1365 CCGTC 1370

RESULT 7  
US-09-796-753-57  
; Sequence 57, Application US/09796753  
; Publication No. US20030027998A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796.753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,094  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 09/474,071  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/474,072  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/514,010  
; PRIOR FILING DATE: 2000-02-25  
; PRIOR APPLICATION NUMBER: 09/516,745  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/572,002  
; PRIOR FILING DATE: 2000-05-14  
; PRIOR APPLICATION NUMBER: 09/597,993  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: 09/599,596  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 09/630,334  
; PRIOR FILING DATE: 2000-07-31

```
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 57
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(652)
; US-09-796-753-57
```

```
Query Match      53.7%; Score 462.2; DB 9; Length 1713;
Best Local Similarity 83.8%; Pred. No. 1.4e-95;
Matches 569; Conservative 0; Mismatches 83; Indels 27; Gaps 3;
```

```
OY 183 GCTGAGCCACTACAAAGCTGTCTGCACGCTCTTCATCAGCCGGACACCCAGGGGGAG 242
Db 1 GCTGAGCCACTACAAAGCTGTCTGCACGCTCTTCATCAGCCGGACACCCAGGGGGAG 60
OY 243 GAGAGAAAGATCCCTGTCTCTGAGAGATGCTGTGGCCCTCGAGAGACAGTGTGAG 302
Db 61 GAGAGAAAGATGATGTCTCTGAGAGATGCTGTGGCCCTCGAGAGATGATGTGAG 120
OY 303 CAAGGAATCCCAAGCGGAGGCTGTACGCGCCCTGCGGCCACAGCCGCTGGGCGT 362
Db 121 ---TGAATGCGGAGGAGGAGGCTGTACGCGCCCTGCGGCCACAGTGTGAG 177
OY 363 GCGGCGCTTGGCCGAGCGGAGGAGGCTGTACGCGCTTCAGCCGCCACCTATCCGTAAC 422
Db 178 GCGGCGCTTGTCCAGGCGG-----AGCCGATTCACACCCACTACCTACCTGCA 228
OY 423 GCAGGAATGAGCTGTCCACCCACCATCTGTGCTGTGAGAGGGGAGAGGCCACCTTA 482
Db 229 GCAGGAATGAGCTGTCCACCCACCATCTGTGCTGTGAGAGGGGAGAGGCCACCTTA 288
OY 483 CCAAGGCGCTTGCACCTCTGAGCTGTGGGAGCCCGAGACAGCTGGAATGTAACGGGA 542
Db 289 CCAAGGCGCTTGCACCTCTGAGCTGTGGGAGCCCGAGACAGCTGGAATGTAACGGGA 348
OY 543 GTGCGTCCGCGCAACCCCAAGAACATCTTGACAGTGTGATGATGATGATGATGATG 602
Db 349 ATCTGTGCGCGCAACCCCTTAACCGAGCATCTTGAGAGTGTGATGATGATGATGATG 408
OY 603 GCTGGGCGGCGCCGCGCCCGCCAGCAGTAATCTGGGATCAGCGCCACGCTTACGGGAG 662
Db 409 GCTGGGCGGCGCCCTGTGTCGCCCGCAGTAATCTGGGATCAGCGCCACGCTTACGGGAG 468
OY 663 CCGCGGCGGCGATGAGAGGGGCGCGCCGCACTACAGCAGAGTATGGGCACTACCGGG 722
Db 469 CCGTGGGCGGATGAGAGGGGCGCGCCGCACTACAGCAGAGTATGGGCACTACCGGG 528
OY 723 GTCTCTCTTCCAGCAGCAGAGAGAGTGTGGCGCCCTCTTGTGTGAGAGGGAGCCGGCT 782
Db 529 GTCTCTCTTCCAGCAGCAGAGAGAGTGTGGCGCCCTCTTGTGTGAGAGGGAGCCGGCT 588
OY 783 CCACGACACACATGCGCGCCCTTACAGAGCGAGCATCTGGAGCAAGAGAAAGATTA 842
Db 589 CCATCACTGTGACATGTGCGCCACTGGA-----GAACAAGGAGAGAGAGAA 633
OY 843 ACAGAAAGAGACACCTCTC 861
Db 634 ACAGAAAGGTCAACCCCTC 652
```

RESULT 8  
US-09-934-249-14/c

```
; Sequence 14, Application US/09934249
; Patent No. US20020115081A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Richard T.
; APPLICANT: Landshultz, Katherine T.
; APPLICANT: Turtl, Thomas G.
; APPLICANT: Thompson, John F.
; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; FILE REFERENCE: P0738/7001/EP/XA
; CURRENT APPLICATION NUMBER: US/09/934,249
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (639)...(639)
; OTHER INFORMATION: a, c, g, or t/u
; US-09-934-249-14
```

```
Query Match      51.8%; Score 445.6; DB 10; Length 693;
Best Local Similarity 97.2%; Pred. No. 6.3e-92;
Matches 485; Conservative 0; Mismatches 10; Indels 4; Gaps 3;
```

```
OY 364 CCGCCCTTGGCCAGCGGAGGAGGCTTCCACGCGCTTCAGCCGCTATCCGTAACCTGAG 423
Db 693 CCGCCCTTGGCCAGCGGAGGAGGCTTCCACGCGCTTCAGCCGCTATCCGTAACCTGAG 634
OY 424 CACGAGATGACCTGTCCACCCACCATCTGTGCTGTGAGAGGGGAGAGGCCACCTAC 483
Db 633 CACGAGATGAG-CTGCGCGCCGACCATCTGTGCTGTGAGAGGGG--AGAGGCCACCTAC 577
OY 484 CAGGCGCCCTGTACCTCTGAGTGTGGAGCCCGAGCAGCAGCTGTGAATGAACCGGAG 543
Db 576 CAGGCGCCCTGTACCTCTGAGTGTGGAGCCCGAGCAGCAGCTGTGAATGAACCGGAG 517
OY 544 TCGGTGCGCGCACCCCAAGAACATCTTGCAGAGTACGTGAGTGAATGATGATGATG 603
Db 516 TCGGTGCGCGCACCCCAAGAACATCTTGCAGAGTGAATGATGATGATGATGATG 457
OY 604 CTGGGCGGCGCCCTGCGCCCGCCAGCAGTAACTGTGGGATCAGCGCCAGCTGTACGGCAGC 663
Db 456 CTGGGCGGCGCCCTGCGCCCGCCAGCAGTAACTGTGGGATCAGCGCCAGCTGTACGGCAGC 397
OY 664 GCGGCGGCGATGAGAGGGGCGCGCCGCACTACAGCAGAGTATGGGCACTACCGGG 723
Db 396 GCGGCGGCGATGAGAGGGGCGCGCCGCACTACAGCAGAGTATGGGCACTACCGGG 337
OY 724 TCCTCTCTTCCAGCAGCAGAGAGAGTGTGGCGCCCTCTTGTGTGAG--GGGAGCCGGCT 782
Db 336 TCCTCTCTTCCAGCAGCAGAGAGAGTGTGGCGCCCTCTTGTGTGAGAGGGGAGCCGGCT 277
OY 783 CCACGACACACATGCGCGCCCTTACAGAGCGCAGCCATCTGGAGCAAGAGAGATTA 842
Db 276 CCGCCGACACACATGCGCGCCCTTACAGAGCGCAGCCATCTGGAGCAAGAGAGATTA 217
OY 843 ACAGAAAGAGACCTCTC 861
Db 216 ACAGAAAGAGACCTCTC 198
```

RESULT 9  
US-09-934-249-16  
; Sequence 16, Application US/09934249  
; Patent No. US20020115081A1  
; GENERAL INFORMATION:



```

? TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
? FILE REFERENCE: PO-16.2c1
? CURRENT APPLICATION NUMBER: US/09/783,590
? CURRENT FILING DATE: 2000-02-15
? PRIOR APPLICATION NUMBER: 08/420,856
? PRIOR FILING DATE: 1995-04-12
? PRIOR APPLICATION NUMBER: 08/346,731
? PRIOR FILING DATE: 1994-11-21
? NUMBER OF SEQ ID NOS: 12485
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 3488
? LENGTH: 65
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (51)
? OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-3488

Query Match          6.9%; Score 59.2; DB 10; Length 65;
Best Local Similarity 93.8%; Pred. No. 7.7e-05;
Matches 61; Conservative 0; Mismatches 4; Indels 0; Gaps 0

OY      715 TACCCGGGTCCTCCTCCAGCACGACGAGCAGTGTGGGCCCGCTTCGTGGAGGG 774
Db       1 TACC GGGGTCTCTCTCCAGCACGACGAGCAGTGTGGGCCCGCTTCGTGGAGGA 60
OY      775 ACCCG 779
Db       61 CCCGG 65

RESULT 13
US-10-184-644-332/C
? Sequence 332, Application US/10184644
? Publication No. US20030044930A1
? GENERAL INFORMATION:
? APPLICANT: Baker, Kevin P.
? APPLICANT: Chen, Jian
? APPLICANT: Desnoyers, Luc
? APPLICANT: Goddard, Audrey
? APPLICANT: Godowski, Paul J.
? APPLICANT: Gurney, Austin L.
? APPLICANT: Pan, James
? APPLICANT: Smith, Victoria
? APPLICANT: Watanabe, Colin K.
? APPLICANT: Wood, William I.
? APPLICANT: Zhang, Zemin
? TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
? FILE REFERENCE: P3430R1C227
? CURRENT APPLICATION NUMBER: US/10/184,644
? CURRENT FILING DATE: 2002-06-28
? Prior Application removed - See File Wrapper or Palm
? NUMBER OF SEQ ID NOS: 612
? SEQ ID NO 332
? LENGTH: 520
? TYPE: PRT
? ORGANISM: Homo Sapien
US-10-184-644-332

Query Match          5.3%; Score 46; DB 9; Length 520;
Best Local Similarity 20.2%; Pred. No. 0.11;
Matches 73; Conservative 79; Mismatches 210; Indels 0; Gaps 0;

OY      468 GGAGCCCCCACCATTACCAAGGGGCCCTGCACCTCCAGCTTGAGGAGCCCCGAGCAGACT 527
Db       508 GHDDCWMSWASGW.AS.ACYG.BNH.W..CACTCBMR.TYCMSRG.KYG.B.THSN.W.H 449
OY      528 GGAACGTGAACCGGGAGTGTGTGGCGGCACCCCCAAACAGAACCATCTTCGACAGTGACCT 587
Db       448 HG.ACWAASRRB.TYCYNSSCR.YBSBN.CV.CM.CMRC.CSCCMCR.HCM.CT.C.T.389

```





Tue Mar 18 08:50:01 2003

us-09-934-249-3.rnpb

Page 10

OY 748 AGTGGGCGCCCTCTTGTGGAGGGGACCCGGCTCCACCACACACATCGGGCCCTA 807  
Db 496 ACCCAGCGCATCCGCGCCCTGTGTGAGACCCAGCTGTGTGACGCGACGCTGGCGAG 555  
OY 808 GAGAGCGCAGCCATCTGGAGCAA 830  
Db 556 GAGGCGGTGTGTATCCGACCGA 578

Search completed: March 17, 2003, 18:33:25  
Job time : 77.3941 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 11:34:54 ; Search time 1617.04 Seconds

(without alignments)  
8623.366 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861

Sequence: 1 atgaccgcgttgatgggggt.....aacagaagacaccccttc 861

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Database :

EST:\*

1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	811.8	94.3	1007	9	AL578575 AL578575
C 2	792.8	92.1	949	9	AL517150 AL517150
C 3	743.6	86.4	967	14	BQ641849 BQ641849
C 4	700	81.3	1046	14	BM922276 BM922276
C 5	678.6	78.8	805	9	AL558881 AL558881
C 6	667.4	77.5	916	14	BQ954555 BQ954555

7	658.6	76.5	898	9	AL558882	AL558882
8	614.6	71.4	1079	11	BC023092	BC023092
9	607.4	70.5	609	14	BQ636742	BQ636742
10	603.6	70.1	1207	11	AK008976	AK008976
11	578.4	67.2	890	14	BQ690750	BQ690750
12	568.2	66.0	782	14	BQ015170	BQ015170
13	550	63.9	551	13	BM141979	BM141979
14	529.8	61.5	894	13	BI851941	BI851941
15	519.4	60.3	729	14	BQ575741	BQ575741
16	516.8	60.0	730	14	BM677602	BM677602
17	501.6	58.3	588	13	BM483503	BM483503
18	468	54.4	857	12	BG323347	BG323347
19	461	53.5	763	13	BI646175	BI646175
20	445.6	51.8	693	9	AT761441	AT761441
21	445.4	51.7	655	14	BQ681705	BQ681705
22	445.4	51.7	1280	14	BQ691500	BQ691500
23	438.6	50.9	825	9	AL543170	AL543170
24	430	49.9	841	13	BI156703	BI156703
25	428.8	49.8	974	10	BB624904	BB624904
26	428.2	49.7	655	13	BI853324	BI853324
27	416.2	48.3	587	13	BI083462	BI083462
28	411	47.7	582	10	BE553323	BE553323
29	403	46.8	467	10	BE666930	BE666930
30	401.6	46.6	900	12	BF036086	BF036086
31	400	46.5	626	14	BM974296	BM974296
32	393	45.6	461	14	BM712680	BM712680
33	392.2	45.6	990	12	BG675643	BG675643
34	392	45.5	464	12	BF026695	BF026695
35	389.8	45.3	633	14	BM714472	BM714472
36	388	45.1	436	13	BM482193	BM482193
37	375.6	43.6	931	12	BF161499	BF161499
38	374	43.4	844	14	BQ686793	BQ686793
39	369.6	42.9	641	10	AM071693	AM071693
40	369	42.9	559	12	BE855409	BE855409
41	366.2	42.5	581	12	BG088400	BG088400
42	365.6	42.5	577	12	BG075859	BG075859
43	360.8	41.9	454	9	AA027926	AA027926
44	359.4	41.7	674	9	AI972096	AI972096
45	358.8	41.7	570	14	BQ575582	BQ575582

## ALIGNMENTS

RESULT 1  
AL578575/c 1007 bp mRNA linear EST 16-FEB-2001  
LOCUS AL578575 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0DK001YC24 3  
DEFINITION prime, mRNA sequence.  
ACCESSION AL578575  
VERSION AL578575.1 GI:12942781  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1007)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.  
FEATURES  
source location/Qualifiers  
1..1007  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DK001YC24"  
/clone\_id="LTI\_NFL006.PL2"  
/tissue\_type="placenta"  
/note="Vector: PCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end



QY 404 CCACCTATCCGTACCTGACAGACAGATGACCTGACACCCACCATCTGCTGACAGC 463  
|||||  
Db 590 CCACCTATCCGTACCTGACAGACAGATGACCTGACACCCACCATCTGCTGACAGC 531  
QY 464 GGGAG 523  
|||||  
Db 530 GGGAG 471  
QY 524 AGCTGGAATGAAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 583  
|||||  
Db 470 AGCTGGAATGAAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 411  
QY 584 ACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 643  
|||||  
Db 410 ACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351  
QY 644 GCGCAG 703  
|||||  
QY 350 GCGCAG 291  
Db 704 TCATGAG 231  
|||||  
QY 290 TCATGAG 231  
|||||  
Db 764 TGCTGAG 823  
|||||  
QY 230 TGCTGAG 111  
Db 824 GGAGCAAG 861  
QY 170 GGAGCAAG 133

RESULT 3  
LOCUS B0641849 967 bp mRNA linear EST 15-JUL-2002  
DEFINITION AGENCOURT 8287174 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:6292265  
5', mRNA sequence.  
ACCESSION B0641849  
VERSION B0641849.1 GI:21766021  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 967)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM2493 row: 9 column: 18  
High quality sequence stop: 571.  
Location/Qualifiers  
1. '967  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6292265"  
/clone\_lib="NIH\_MGC\_43"  
/tissue\_type="normal pigmented retinal epithelium"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(c). Library constructed by Ling Hong

## FEATURES

source

in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH-MGC Library. |"

BASE COUNT 194 a 334 c 296 g 143 t  
ORIGIN  
Query Match 86.4%; Score 743.6; DB 14; Length 967;  
Best Local Similarity 99.5%; Pred. No. 1.1e-150;  
Matches 746; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 112 GACCTGGAGAGTTTTCATGATCATCATCATGATGATGATGATGATGATGATGATGATG 171  
Db 1 GACCTGGAGAGTTTTCATGATCATCATCATGATGATGATGATGATGATGATGATGATG 60  
QY 172 ATACAGTCCCTGCTGATGACCACTCAAGCTGTGTGACAGGTCTTTCATCAGCCGACAGC 221  
Db 61 ATACAGTCCCTGCTGATGACCACTCAAGCTGTGTGACAGGTCTTTCATCAGCCGACAGC 120  
QY 232 CAGGGGGGAG 291  
Db 121 CAGGGGGGAG 180  
QY 292 ACAAGTGTACAGCAACGGAATCCAGAGCCGACAGGTCTACGCCGCCCTCGGCCACCGAC 351  
Db 181 ACAAGTGTACAGCAACGGAATCCAGAGCCGACAGGTCTACGCCGCCCTCGGCCACCGAC 240  
QY 352 CGCCTGGCGTGGCGCCCTTTCGCGCCAGCGGAGAGCGCTTCCAGCGCTTTCAGGCCACTAT 411  
Db 241 CGCCTGGCGTGGCGCCCTTTCGCGCCAGCGGAGAGCGCTTTCAGCGCTTTCAGGCCACTAT 300  
QY 412 CGGTACCTGACAGACAGAGATGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 471  
Db 301 CGGTACCTGACAGACAGAGATGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 360  
QY 472 CCCCCACCTTACAG 531  
Db 361 CCCCCACCTTACAG 420  
QY 532 CTGACAGCGGAG 591  
Db 421 CTGACAGCGGAG 480  
QY 592 GATAGTGCAG 651  
Db 481 GATAGTGCAG 540  
QY 652 TGCTAGCGAG 711  
Db 541 TGCTAGCGAG 600  
QY 712 CATTACCCCGGGGTCCTTCTTCCAGACACACAGACAGAGAGAGAGAGAGAGAGAG 771  
Db 601 CATTACCCCGGGGTCCTTCTTCCAGACACACAGAGAGAGAGAGAGAGAGAGAGAG 660  
QY 772 GGGAGCCGGGTCTTCCAGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 831  
Db 661 GGGAGCCGGGTCTTCCAGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
QY 832 GAGAGAGATTAACAGAAAGAGACACCTCTC 861  
Db 721 GAGAGAGATTAACAGAAAGAGACACCTCTC 750  
RESULT 4  
LOCUS BM922276 1046 bp mRNA linear EST 12-MAR-2002  
DEFINITION AGENCOURT 6707077 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5754437  
5', mRNA sequence.  
ACCESSION BM922276  
VERSION BM922276.1 GI:19372655  
KEYWORDS EST.  
SOURCE human.



```

Db 805 GTGCTCTGTGAGCCACTACAGCTGTCTGACGGTCTCTTCATCAGCCGGAACAGCCAGG 746
Qy 237 GCGGAGGAGAGATGATGCTGTCTCAGAGATGCTGTGGGCTTGAGGAGACAGT 296
Db 745 GCGGAGGAGAGATGATGCTGTCTCAGAGATGCTGTGGGCTTGAGGAGACAGT 686
Qy 297 GTCAGGACAGGATGCCAGAGCCGAGGTCGTACGCCCCGCTGCGCCAGCAGCCGCT 356
Db 685 GTCAGGACAGGATGCCAGAGCCGAGGTCGTACGCCCCGCTGCGCCAGCAGCCGCT 626
Qy 357 GCGGCTCGCCGCTGCGCCAGAGCCGCTTCACCGCTTCACGCCACTATCCGTA 416
Db 625 GCGGCTCGCCGCTGCGCCAGAGCCGCTTCACCGCTTCACGCCACTATCCGTA 566
Qy 417 CCGTCACAGAGATGACCGCCACCATCTGCTGTGACAGCGGGAGAGCCGCC 476
Db 565 CCGTCACAGAGATGACCGCCACCATCTGCTGTGACAGCGGGAGAGCCGCC 506
Qy 477 ACCCTACAGAGCCCTGACACCTTCAGCTTCGGGAGCCCGAGCAGCAGTGAAGTAA 536
Db 505 ACCCTACAGAGCCCTGACACCTTCAGCTTCGGGAGCCCGAGCAGCAGTGAAGTAA 446
Qy 537 CCGGAGTCTGCTGCGGACCCCAACAGAACATCTTCAGACAGTATGATAG 596
Db 445 CCGGAGTCTGCTGCGGACCCCAACAGAACATCTTCAGACAGTATGATAG 386
Qy 597 TGCCAGCTGGGGGCGCCCTGCGCCAGAGTACCTGGGCGATCGGCGCAGTGTGTA 656
Db 385 TGCCAGCTGGGGGCGCCCTGCGCCAGAGTACCTGGGCGATCGGCGCAGTGTGTA 326
Qy 657 CCGGAGCGGGGCGATGAGAGGGGCGCGCCACCTACAGCAGAGTATCGGCCACTA 716
Db 325 CCGGAGCGGGGCGATGAGAGGGGCGCGCCACCTACAGCAGAGTATCGGCCACTA 266
Qy 717 CCGGAGGCTCTCTCTTCAGACACAGCAGAGTGGGCGCTCTCTCTGAGGGGAC 776
Db 265 CCGGAGGCTCTCTCTTCAGACACAGCAGAGTGGGCGCTCTCTCTGAGGGGAC 206
Qy 777 CCGGCTCCACACACACATCGCGCCCTAGAGAGGCGCATCTGGAGCAAGAGAA 836
Db 205 CCGGCTCCACACACACATCGCGCCCTAGAGAGGCGCATCTGGAGCAAGAGAA 146
Qy 837 GGATTAACAGAAAGACACCTCTC 861
Db 145 GGATTAACAGAAAGACACCTCTC 121

```

RESULT 6  
2954555

DEFINITION AGENCOURT\_8825282\_lupski\_sciatic\_nerve\_Homo\_sapiens\_cDNA\_clone  
LOCUS IMAGE:6204609 5', mRNA sequence.

ACCESSION B0934555  
VERSION B0934555.1 GI:22370033  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 916)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-f@mail.nih.gov](mailto:cgaps-f@mail.nih.gov)  
Tissue Procurement: Dr. James R. Lupski  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
[found through the I.M.A.G.E. Consortium/LNL at:](http://image.lnl.gov)  
<http://image.lnl.gov>  
Plate: L1AM13626 row: c column: 10

```

FEATURES             High quality sequence stop: 669.
source               Location/Qualifiers
                    1..916
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:6204609"
                    /clone_1id="lupski_sciatic_nerve"
                    /sex="male"
                    /tissue_type="sciatic nerve"
                    /dev_stage="adult, 70 yr"
                    /lab_host="DH10B"
                    /note="Vector: PCMV-SPORE6 (Life Technologies); Site_1:
                    NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                    directionally cloned using the following adaptors:
                    5'-TCCACCACCGCCGCC-3' and
                    5'-GACTAGTTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >
                    1 kb for average insert length 1.87 kb. This is a primary
                    library, non-amplified. Library constructed by Life
                    Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                    College of Medicine) and is available through Life
                    Technologies."
BASE COUNT          162 a 354 c 273 g 127 t
ORIGIN
Query Match          77.5%; Score 667.4; DB 14; Length 916;
Best Local Similarity 95.3%; Pred. No. 3.2e-134;
Matches 731; Conservative 0; Mismatches 31; Indels 5; Gaps 4;
Qy 1 ATGCACCGCTGATGAGGGGTACAGACACCGCCGCCGCCGCGGACCCCAATGTC 60
Db 70 ATGCACCGCTGATGAGGGGTACAGACACCGCCGCCGCCGCGGACCCCAATGTC 129
Qy 61 TCCGCAAGTCGCACTGCAAAACGCTTTGTTCCAGAGCATGAGATCAGGAGCTGGAG 120
Db 130 TCCGCAAGTCGCACTGCAAAACGCTTTGTTCCAGAGCATGAGATCAGGAGCTGGAG 189
Qy 121 TTTGTTGATCATCATCATCATGCTGTGATGATGATGATGATGATGATGATGATGATG 180
Db 190 TTTTTCGATCATCATCATCATGCTGTGATGATGATGATGATGATGATGATGATGATG 249
Qy 181 CTGCTGAGCCACTACAGCTGTGCAAGCGCTCTTCATCAGCCGCGACAGCGGCGG 240
Db 250 CTGCTGAGCCACTACAGCTGTGCAAGCGCTCTTCATCAGCCGCGACAGCGGCGG 309
Qy 241 AGAGAGAAATGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 310 AGAGAGAAATGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 369
Qy 301 GGCAGAGAAATCCAGAGCCGACAGCTACGCGCCGCTGCGCCACAGCAGCGCTGAGC 360
Db 370 GGCAGAGAAATCCAGAGCCGACAGCTACGCGCCGCTGCGCCACAGCAGCGCTGAGC 429
Qy 361 GTGCGGCGCTTCCGCGCCAGCGGAGCGCTTCACCGCTTCAGCGCCACCTATCCGTAC 420
Db 430 GTGCGGCGCTTCCGCGCCAGCGGAGCGCTTCACCGCTTCAGCGCCACCTATCCGTAC 489
Qy 421 CAGACAGAGATCGACCTGCGCCACCAACCATCTGCTGTCAAGCGGGAGAGAGCCGCCACC 480
Db 490 CAGACAGAGATCGACCTGCGCCACCAACCATCTGCTGTCAAGCGGGAGAGAGCCGCCACC 549
Qy 481 TACAGAGGCGCCCTGACCGCTCAGCTTCGGGAGCCCGAGCAGCTGGAAGTGAACCGG 540
Db 550 TACAGAGGCGCCCTGACCGCTCAGCTTCGGGAGCCCGAGCAGCTGGAAGTGAACCGG 609
Qy 541 GAGTCGGTGGCGCACCCCAACAGAACATCTTTCAGACAGTGAAGTGAAGTGAAGTGAAG 600
Db 610 GAGTCGGTGGCGCACCCCAACAGAACATCTTTCAGACAGTGAAGTGAAGTGAAGTGAAG 669
Qy 601 -AGGCTGGGCGGCGCCCGCCCGCCAGCAGTATCGGCGCATGAG -CGCCACGTGCTAGC 658
Db 670 AAGGCTGGGCGGCGCCCGCCCGCCAGCAGTATCGGCGCATGAGCGCCACAGTGTACG 729
Qy 659 GCAGCGGCGGCG -GCATGAGAGGGCGCGCGCCACCTACAGCGCATGATGCGGCA -CT 715

```

RESULT 7	898 bp	mRNA	linear	EST 16- <i>FEb</i> -2001
AL558882				
LOCUS				
DEFINITION	AL558882	LT1.NFL008.TC3	Homo sapiens	cDNA clone CS0D015YF12 5
ACCESSION	AL558882			
VERSION	AL558882.1	GI:12903838		
KEYWORDS	EST,			
SOURCE	human.			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 898)	Li, B., Gruber, C., Jesse, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished (2001)	Contact: Genoscope
				Genoscope - Centre National de Sequencage
				BP 191 91006 EVRY cedex - France
				Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
				Location/Qualifiers
				1. 898

BASE COUNT	ORIGIN
196 a	"http://nulllength.invtirogen.com"
310 c	260 g 126 t 6 others

Best Local Similarity	76.5%	Score 658.6;	DB 9;	Length 898;
Matches 680; Conservative	99.3%;	Pred. No. 2.6e-13;		
		2; Mismatches		

		1; indels	2; gaps
QY	177	GTGCTGTGAGCCACTACAAAGTGTGTGCAAGGGTCCCTTATTATAGCGGGACAGCCAGGG	236
Db	1	GTGCTGTGTAGGCACCTACAAAGCTGTGTGCAAGGTTCTTATTATAGCGGGACAGCCAGGG	60
QY	237	GCAGAGAGAGAAGATGTGCCCTTCTCTCAAGATGTCCTGTGGCCCTGGAGAGACAGT	296
Db	61	GCAGAGAGAGAAGATGTGCCCTTCTCTCAAGATGTCCTGTGGCCCTGGAGAGACAGT	120
QY	297	GTCAGGCAACGGAATCCACAGAGCCGCAAGTCTAGGCCCGCTCGGCCACAGCAAGCCCT	356
Db	121	GTCAGGCAACGGAATCCACAGAGCCGCAAGTCTAGGCCCGCTCGGCCACAGCAAGCCCT	179
QY	357	GGCCGTGTGCCCTTCGCCCCAGGCGGAGCCCTTCACCCCTTCACAGCCACATATCCGTA	416
Db	180	GGCCGTGTGCCCTTCGCCCCAGGCGGAGCCCTTCACCCCTTCACAGCCACATATCCGTA	239
QY	417	CCGACGACGACGATCGACTGTGCACCCACACATCTCGGTGTCAAGAGGGAGAGGCCGCC	476
Db	240	CCGACGACGACGATCGACTGTGCACCCACACATCTCGGTGTCAAGAGGGAGAGGCCGCC	299

OY	477	ACCCATACAGGGGCCCTTCGACACCTCCAGCTTGCGGAGACCCCGAGCAGCATGTGGAACCTGAA	536
Db	300	ACCCATACAGGGGCCCTTCGACACCTCCAGCTTGCGGAGACCCCGAGCAGCATGTGGAACCTGAA	536
OY	537	CCGGGAGCTCGGTGCGGCGACACCCCAACAGAACCACTCTTGACAGGTACTGATGATAG	359
Db	360	CCGGGAGCTGTGTGTGCGGCGACACCCCAACAGAACCACTCTTGACAGGTACTGATGATAG	359
OY	597	TGCGAAGCTGTGGCGGGGCCCTTGCCCCCCCCAGCATTAACCTCGGGCATCAGCGCCACGTGTA	419
Db	420	TGCGAAGCTGTGGCGGGGCCCTTGCCCCCCCCAGCATTAACCTCGGGCATCAGCGCCACGTGTA	419
OY	657	CGGCAACGGCGGGGGCCATATGAGAGGGGGCCGCCACCTATAGCAGGAGTATGTGGGCACCTA	479
Db	480	CGGCAACGGCGGGGGCCATATGAGAGGGGGCCGCCACCTATAGCAGGAGTATGTGGGCACCTA	479
OY	717	CGCGGGGTCTCTCTTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	539
Db	540	CGCGGGGTCTCTCTTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	539
OY	777	CGCGGTCCACACACACACATGCGGGCCCCCAAGAGACGACGACGACGACGACGACGACGACGACGAC	599
Db	600	CGCGGTCCACACACACACATGCGGGCCCCCAAGAGACGACGACGACGACGACGACGACGACGACGAC	599
OY	837	GGATTAACACAAAGAGACACCTCTC 861	658
Db	659	GGATTAACACAAAGAGACACCTCTC 863	658

RESULT 8	
BC023092	
LOCUS	
DEFINITION	BC023092 Mus musculus, Nedd4 WW binding protein 4, clone IMAGE:398996,
ACCESSION	Mus musculus, Nedd4 WW binding protein 4, clone IMAGE:398996,
VERSION	BC023092
KEYWORDS	BC023092.1 GI:18605637
SOURCE	HTC.
ORGANISM	house mouse, Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 1079)
AUTHORS	Straussberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (04-FEB-2002) NCBI

REMARK COMMENT
<p>NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>  Contact: MGC help desk</p>

CDNA Library Construction: Gilbert Smith, Ph.D.  
 CNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LINT)  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: amg@bcm.tmc.edu  
 Guinardine, P. H., Garcia, A.M., Lu, X., Huijck, S.W., Hale, S.M.,  
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>  
 Series: IRMA Plate: 36 Row: 6 Column: 6  
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein  
 This clone has the following problem: frame shifted.  
 Location/Qualifiers

/organism="Mus musculus"



```

/db_xref="taxon:10090"
/map="CECH II"
/clone="IMAGE:3989996"
/tissue_type="Mammary tumor metastatized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clone) outgrowth infected with the virus MMTV."
/clone_id="NCI_CGAP_Lu29"
/lab_host="DH10B"
/Note="Vector: pCMV-Sport6"
BASE COUNT      271 a      341 c      289 g      178 t
ORIGIN

```

```

Query Match      71.4%; Score 614.6; DB 11; Length 1079;
Best Local Similarity 86.8%; Pred. No. 8.6e-123;
Matches 702; Conservative 0; Mismatches 95; Indels 12; Gaps 2;

```

```

QY 5 ACCGCTGATGAGGGGTCAACAGCACCGCCGCCGCCGCCGCCGCCGCCCAATGTCCT 64
DB 15 ACCGCTGATGAGGGGTCAACAGCACCGCCGCCGCCGCCGCCGCCCAATGTCCT 74
QY 65 GCAGGTGCACTGCAACGGCTTGTGTTCCAGAGCATGAGATGACGAGGAGTTTG 124
DB 75 GCGGTGCACTGCAACGGCTTGTGTTCCAGAGCATGAGATGACGAGGAGTTTG 134
QY 125 TTCAGATCATCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 184
DB 135 TGCAATCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 194
QY 185 TGACCACTACAGCTGTCTGACAGGTCCTCATACAGCCGACACAGCCAGGCGAGGA 244
DB 195 TGACCACTACAGCTGTCTGACAGGTCCTCATACAGCCGACACAGCCAGGCGAGGA 254
QY 245 GAGAGAGTGGCCCTGCTCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304
DB 255 GAGAGAGTGGCCCTGCTCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
QY 305 ACGGATGCCAGAGCCGACAGGTCTACGCGCCGCTGCGCCACCGACCGCTGCGCTG 364
DB 313 -TGGAAATGCCAGAGCCGACAGGTCTACGCGCCGCTGCGCCACCGACCGCTGCGCTG 371
QY 365 CGCCCTTGCCAGAGGAGGAGCGCTTCCAGCGCTTCCAGCGCTTCCAGCGCTTCCAG 424
DB 372 CCCCCTTATCAGCAGCG-----AGCGATTCACCAACCACTACCCCTACCTGACAG 422
QY 425 ACGGATGCCAGAGCCGACAGGTCTACGCGCCGCTGCGCCACCGACCGCTGCGCTG 484
DB 423 ACGAATTTGCCCTGCAACCGACCATCTCTGATGAGGAGAGAGCCGCCACCTTACC 482
QY 485 AGGCGCCCTGACAGCTTCCAGCTTGGAGACCCGAGCAGACAGCTGAACTGAACCGGAGT 544
DB 483 AGGCGCCCTGACAGCTTCCAGCTTGGAGACCCGAGCAGACAGCTGAACTGAACCGGAGT 542
QY 545 CGGTGCGGACCGCCCAACAGAACATCTTGCAGAGTACGACCTGATGATGATGATG 604
DB 543 CTGTGCGGACCGCCCAACAGAACATCTTGCAGAGTACGACCTGATGATGATGATG 602
QY 605 TGGGCGGCGCCCTGCGCCCGAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 664
DB 603 TGGGCGGCGCCCTGCGCCCGAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 662
QY 665 GCGGCGCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 724
DB 663 GTGGGCGCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
QY 725 CTTCTTTCAGACACAGAGAGAGAGTGGGCGCCCTCTTCTGTGAGGAGGAGGAGGAG 784
DB 723 CTTCTTTCAGACACAGAGAGAGTGGGCGCCCTCTTCTGTGAGGAGGAGGAGGAGGAG 782
QY 785 ACGACACAGACATGCGCGCCCTAGAGAGC 813
DB 783 ATCACTGCACATTTGCCCGCCTGAGAGC 811

```

```

RESULT 9
B0636742
LOCUS
DEFINITION
hd13h06.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
ACCESSION
B0636742
VERSION
B0636742.1 GI:21761201
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 609)
Wistow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A., Touchman
J.W., Bouffard, G., Smith, D. and Peterson, K.
Expressed sequence tag analysis of human retina for the NEBank
Project: Rebinding, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 13 row: h column: 06
Seq primer: M13Rpl reverse primer (ABI).
Location/Qualifiers
1. 609
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hd13h06"
/clone_id="Human Retina cDNA (Un-normalized, unamplified)"
): hd/he
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/Note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the Superscript Plasmid System full
details of which are contained in the manufacturer's
instruction manual. (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-GGACTAGTCTAGATGCGGAGCGCGCCGCTT)15-3'
]. EST analysis was performed on the unamplified library
at the NIH Intramural Sequencing Center (NISC)."
BASE COUNT      114 a      238 c      182 g      75 t
ORIGIN
Query Match      70.5%; Score 607.4; DB 14; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.8e-121;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

REFERENCE	TITLE
ADACHI, J., AKIYAMA, K., AKIYAMA, C. ET AL.	FUNCTIONAL ANNOTATION OF A FULL-LENGTH MOUSE CDNA COLLECTION
5 (bases 1 to 1207)	
11217851	
PUBLISHED	JOURNAL MEDICINE
	NUMBER 409 (6821), 685-690 (2001)
	NATURE 409 (6821), 685-690 (2001)
	12185560
	11217851
	ADACHI, J., AKIYAMA, K., AKIYAMA, C. ET AL.

TITLE  
JOURNAL

COMMENT

## FEATURES

Library was prepared and sequenced in Mouse Genome  
Sciences Center and Genome Science Laboratory in Riken  
mouse tissues. First Strand cDNA was primed with a primer  
subsequently enriched for full-length by cap-trapper. cDNA went  
into a second round of normalization to Rot = 5.0 and subtraction Rot  
= 15'.  
Second strand cDNA was prepared with the primer adapter of  
5'-GAGATGTCGACGATTAATTAATTCGCCCCCCCC 3'}. cDNA was cleaved  
with NotI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

```
1. Location/Qualifiers
    /organism="Mus musculus"
    /db_xref="NCBI:60"
    /db_xref="RANTOM.DB:2210418102"
    /db_xref="MGD:MGI:1902457"
    /taxon="taxon:10090"
    /clone="2210418102"
    /sex="male"
    /tissue_type="stomach"
    /clone_id="RIKEN full-length enriched mouse cDNA library"
```

```

CDS
    /dev_stage="adult"
    391..1173
    /note="Nedd4 WW binding protein 4
    data source:MCD, source key:MGI:1929600, evidence:ISS
    putative"
    /codon_start=1
    /protein_id="BAB26001.1"
    /db_xref="GI:12843489"
    /db_xref="MGI:1929600"
    /translation="MSPARATADRSLEPSEITELEPVQIVVIVVMVWVWITCL
    SHTKLSRSTISHSDARRRDELSSBGLWPESEVSGMPPEVYAPPRDRLAV
    PPIQRREQPTPYLQHEIALPPTISLSGEBEPPOGCTQLRPEOOLENRES
    VRAPIPNRTFSDLDISTMLGCPSPSSNGISATCYSSGSRGMPPTYSVYGHVP
    GSSFOHOOSNGPSLSLEGTRLHSHIAPLENKEKXKXKGPL"
BASE COUNT      200 a      459 c      374 g      173 t      1 others
ORIGIN
Query Match      70.1%; Score 603.6; DB 11; Length 1207;
Best Local Similarity 85.2%; Pred. No. 2.1e-120;
Matches 731; Conservative 0; Mismatches 99; Indels 28; Gaps 4;
QY 5 ACCGCTGATGAGGGGTCAACAGCACCGCCGCCGCCGCCGCCGCAATGTCCT 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 340 ACCGCTGATGAGGGGTCAACAGCACCGCCGCCGCCGCCGCCGCAATGTCCT 399
QY 65 GCAGTCACTG-CAAACGCTTTGTTCCAGAGCAGTGGAGTACAGGAGGAGTT 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 400 GCGGTGCACTGCCAGCGCTCTTTGTTCCAGCAGTGGAGTACAGGAGGAGTT 459
QY 124 GTTCAGATCATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 460 GTTCAGATCATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519
QY 184 CTGAGCCATCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 520 CTGAGCCATCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579
QY 244 AGAAGATGCGCTGTCTCTGAGAGATGCTGTGTGTGTGTGTGTGTGTGTGT 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 580 AGAAGATGCGCTGTCTCTGAGAGATGCTGTGTGTGTGTGTGTGTGTGTGT 638
QY 304 AACGGAATCCAGAGCGGAGGTTACGCGCCGCCGCCGCCGCCGCCGCCGCTG 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 639 --TGAATGCGGAGCAGAGCTATATGCCGCCGCCGCCGCCGCCGCCGCCGCTG 696
QY 364 CCGGCTTCCGCGGAGGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGG 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 697 CCGGCTTCCAGGAGGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGG 747
QY 424 CAGAGATGAGCTGCGACCCAGCATCTCGGTTCAGAGCGGAGGAGGAGGAGGAG 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 748 CAGAGATGAGCTGCGACCCAGCATCTCGGTTCAGAGCGGAGGAGGAGGAGGAG 807
QY 484 CAGGCGCCCTGACCCCTCGAGCTTCGGAGCCCGAGCAGCAGTGAAGTGAACCG 543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 808 CAGGCGCCCTGACCCCTCGAGCTTCGGAGCCCGAGCAGCAGTGAAGTGAACCG 867
QY 544 TCGGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 868 TCTGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 927
QY 604 CTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 928 CTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 987
QY 664 GCGGCGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 988 GTGGGCGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1047
QY 724 TCTGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 783
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1048 TCTGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1107

```

```

QY 784 CACCACACATCGCGCCCTGAGAGAGCGACGACCATCTGAGCAAAAGAGATAAA 843
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1108 CATCACTGCACTTGTCCCACTGGA-----GAACAGAGAGAGAGAGAAA 1152
QY 844 CAGAAAGACACCTCTC 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1153 CAGAAAGTCAACCCCTC 1170
RESULT 11
B0690750      890 bp      mRNA      linear      EST 15-JUL-2002
LOCUS      AGENCOURT 8046394 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209341
DEFINITION      5', mRNA sequence.
ACCESSION      B0690750
VERSION      B0690750.1 GI:21816066
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 890)
AUTHORS      NIH-MGC http://mgc.ncl.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgaps-remail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: L1CM2368 row: h column: 14
            High quality sequence stop: 627.
FEATURES
    source
        1..890
            location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_id="NIH_MGC_110"
            /tissue_type="ductal carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: pancreas; Vector: pOT8; Site_1: XhoI;
            Site_2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCACGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit.
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."
BASE COUNT      168 a      303 c      273 g      146 t
ORIGIN
Query Match      67.2%; Score 578.4; DB 14; Length 890;
Best Local Similarity 93.5%; Pred. No. 5.0e-115;
Matches 615; Conservative 0; Mismatches 41; Indels 2; Gaps 1;
QY 87 TTGTTCAGAGATCAGAGATCAGAGATCAGAGATCAGAGATCAGAGATCAGAG 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 105 TCTCTCGAAGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 164
QY 147 GGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 165 GGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 224
QY 207 ACGGTCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 225 ACGGTCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 284
QY 267 AGGATGCTGTGCGCTCGAGAGACAGTGTCAAGGAGGAGGAGGAGGAGGAGGAG 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 285 AGATGCTGTGGCCCTCGAGAGACACAGTGTACGAGCAAGCAATCCAGAGCCGACAGT 344
QY 327 CTAGCCCGCCCTCGGCGCCACGACCGCTGCGCTGCGCCCTTCGCGCCAGCGGAGCG 386
Db 345 CTAGCGCCCGCCCTCGGCGCCACGACCGCTGCGCTGCGCCCTTCGCGCCAGCGGAGCG 404
QY 387 CTTCACCGCTTCACCGCCACCTATCCGCTACCTGAGCAGAGATGCACTGCGCCAC 446
Db 405 CTTCACCGCTTCACCGCCACCTATCCGCTACCTGAGCAGAGATGCACTGCGCCAC 464
QY 447 CATCTGCTGTGACAGCGGAGGAGCGCCCGCCACCTACAGGCGCCCTGACCCCTCAGCT 506
Db 465 CATCTGCTGTGACAGCGGAGGAGCGCCCGCCACCTACAGGCGCCCTGACCCCTCAGCT 524
QY 507 TCGGAGCCCGCCGAGCAGCTGGAATGAAACCGGAGTGTGCGCGGCGACCCCGCAACAG 566
Db 525 TCGGAGCCCGCCGAGCAGCTGGAATGAAACCGGAGTGTGCGCGGCGACCCCGCAACAG 584
QY 567 AACCATCTTGCACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 626
Db 585 AACCATCTTGCACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 644
QY 627 CAGTAATCTGCGGATCTGCGCCGACGCTGCTACGCGAGCGCGCGCGCATGGA--GGGGCGG 684
Db 645 CAGTAATCTGCGGATCTGCGCCGACGCTGCTACGCGAGCGCGCGCGCATGGA--GGGGCGG 704
QY 685 CCGCCGACCTACAGCGAGGTGATGCGGCGACCTACCGGGGTCTCTCTTCGACGACGAGC 742
Db 705 CGCCCGACCTACAGCGAGGTGATGCGGCGACCTACCGGGGTCTCTCTTCGACGAGC 762

```

```

RESULT 12
BQ015170 782 bp mRNA linear EST 26-MAR-2002
LOCUS UI-H-ED1-axw-k-20-0-UI.s1 NCI CGAP_Ed1 Homo sapiens cDNA clone
DEFINITION IMAGE:3834635 3', mRNA sequence.
ACCESSION BQ015170
VERSION BQ015170.1 GI:19740071
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cga@bcrf.org
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/BLNL at: http://image.llnl.gov
Seg primer: M13 FORWARD
POLYA-Yes.

```

# FEATURES

```

source Location/Qualifiers
1..782
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5834635"
/clone_id="NCI CGAP_Ed1"
/tissue_type="Chondrosarcoma"
/deb_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Left Pubic Bone; Vector: p773-Pac
(Pharmacia) with a modified polylinker; Site:1: EcoR I;
Site:2: Not I; NCI-CGAP Ed1 is a normalized cDNA library
containing the following tissue(s): Chondrosarcoma cell
line C5. The library was constructed according to Bonaldo
, Lennon and Soares, Genome Research, 6:791-806, 1996."

```

First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of the first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dri)8 tail. The sequence tag for this library is GCTCAAGGCT.

TAG TISSUE=chondrosarcoma  
TAG\_SEQ=CGTCAAGGCT

BASE COUNT 109 a 223 c 271 g 176 t 3 others

Query Match 66.0%; Score 568.2; DB 14; Length 782;  
Best Local Similarity 99.0%; Pred. No. 8,7e-113;  
Matches 570; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 286 GAGACACAGTGTGACGAGCAAGCAATCCAGAGCCGAGGCTTACCGCCGCTGCGGCC 345
Db 782 GAGACACAGTGTGACGAGCAAGCAATCCAGAGCCGAGGCTTACCGCCGCTGCGGCC 723
QY 346 ACCGACCGCCCTGCGCGCTGCGCGCCCTTCGCGCGGAGGCGTTCACCGCTTCCAGCC 405
Db 722 ACCGACCGCCCTGCGCGCTGCGCGCCCTTCGCGCGGAGGCGTTCACCGCTTCCAGCC 663
QY 406 ACCTATCGTACCTGACGACGAGATGACCTTCCACCCACCATCTGCTGTGACAGCGG 465
Db 662 ACCTATCGTACCTGACGACGAGATGACCTTCCACCCACCATCTGCTGTGACAGCGG 603
QY 466 GAGAGCGCCCGCCCTTACCAAGGCGCCCTGACCTTCCAGCTTCCGAGACCCCGACAGC 525
Db 602 GAGAGCGCCCGCCCTTACCAAGGCGCCCTGACCTTCCAGCTTCCGAGACCCCGACAGC 543
QY 526 CTGGAACCTGAACCGGAGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 585
Db 542 CTGGAACCTGAACCGGAGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 585
QY 586 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 645
Db 482 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 645
QY 482 GCGACGCTGTACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 423
Db 646 GCGACGCTGTACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 705
QY 422 GCGACGCTGTACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 363
QY 706 ATCGGCACTACCGCGGCGCTCTCTTCACGACAGCAGAGAGTGTGCGCGCGCGCGCGCG 765
Db 362 ATCGGCACTACCGCGGCGCTCTCTTCACGACAGCAGAGAGTGTGCGCGCGCGCGCGCG 303
QY 766 CTGAGGAGGAGCGCGCTTCACGACAGCAGAGAGTGTGCGCGCGCGCGCGCGCGCGCG 825
Db 302 CTGAGGAGGAGCGCGCTTCACGACAGCAGAGAGTGTGCGCGCGCGCGCGCGCGCGCG 825
QY 826 AGCAAGAAGAGAGTAAACGAAAGAGACACCTCTTC 861
Db 242 AGCAAGAAGAGAGTAAACGAAAGAGACACCTCTTC 207

```

```

RESULT 13
BQ141979 551 bp mRNA linear EST 12-MAR-2002
LOCUS BQ141979
DEFINITION IEF2a11.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5677341 5' similar to TR:09UD3 Q9UDJ3 D77187.1
; mRNA sequence.
ACCESSION BQ141979
VERSION BQ141979.1 GI:17152046
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 551)

```



```

OY 79 AAACGCTCTTGTTCAGAGCATGAGATCATCGAGCTGAGTTGTCATCATCATC 138
Db 61 CAGCCCTCTTGTTCAGAGCATGAGATCATCGAGCTGAGTTGTCATCATCATC 120
OY 139 ATCGGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Db 121 ATCGGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 198
OY 199 CTGCTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 181 CTGCTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 258
OY 259 TCTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db 241 TCTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 318
OY 319 CCGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 297
Db 298 CCGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 378
OY 379 CCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 357
Db 358 CCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 438
OY 439 CCACCCACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 408
Db 409 CCACCCACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
OY 499 CTCCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 468
Db 469 CTCCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 558
OY 559 CCAACAGAGCATCTTTCAGAGCATGAGATGATGATGATGATGATGATGATGAT 528
Db 529 CCAACAGAGCATCTTTCAGAGCATGAGATGATGATGATGATGATGATGATGAT 616
OY 617 GCGCCCGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588
Db 589 GCGCCCGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 676
OY 677 AGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 648
Db 649 AGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 736
OY 737 ACCAGAGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 707
Db 708 ACCAGAGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 795
OY 796 ATCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 767
Db 768 ATCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 834

```

```

RESULT 15
LOCUS B0575741/c 729 bp mRNA linear EST 19-JUN-2002
DEFINITION UT-H-E21-bbg-h-14-0-UI.s1 NCI CGAP-Ch2 Homo sapiens cDNA clone
ACCESSION B0575741
VERSION B0575741.1 GI:21479058
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 729)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

```

Tissue Procurement: Dr. Steven Gitellis/ Rush Presbyterian, Dept. of Orthopedics  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu  
 Seq primer: M13 FORWARD  
 POLY-A-yes.

## FEATURES

SOURCE

Location/Qualifiers

1..729

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-H-E21-bbg-h-14-0-UI"

/clone\_lib="NCI-CGAP-Ch2"

/tissue\_type="Chondrosarcoma Grade II"

/dev\_stage="Adult"

/lab\_host="DH10B (Life Technologies)"

/note="Organ: Left Pelvis; Vector: pRT3-Pac (Pharmacia)

with a modified polylinker: Site 1: EcoR I; Site 2: Not I;

following tissue(s): Chondrosarcoma Grade II. The library

was constructed according to Bonaldo, Lennon and Soares,

synthesis was primed with an oligo-dT primer containing a

adaptor, digested with Not I, and cloned directionally

into pRT3-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

(dT)18 tail. The sequence tag for this library is

TGATACACT.

TAG-LIB=UI-H-E21

TAG-TISSUE=grade-2-chondrosarcoma

TAG\_SEQ=ATCTAATG

ORIGIN

101 a 213 c 253 g 162 t

## Query Match

Best local similarity 60.38; Score 519.4; DB 14; Length 729;

Matches 520; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 341 GCGCCCGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
Db 729 GCGCCCGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
OY 401 AGCCACCTATCGCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
Db 669 AGCCACCTATCGCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 610
OY 461 ACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 520
Db 609 ACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 550
OY 521 AGCAGCTGGAAGTGAACCGGGAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
Db 549 AGCAGCTGGAAGTGAACCGGGAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 490
OY 581 GTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
Db 489 GTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 700
OY 641 TCAGGCGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 430
Db 429 TCAGGCGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 370
OY 701 AGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 760
Db 369 AGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 310
OY 761 CTTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 820
Db 309 CTTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 250

```

Tue Mär 18 08:50:02 2003

us-09-934-249-3.rst

Page 13

Oy	821 TCTGGAGCAAAAGAGAAGATTAACAGAAAAAGCACCCCTCTC	861 
Dd	249 TCCTGAGCAAAGAGAAGATTAAACAATAAAGCACCCTCTC	209 

Search completed: March 17, 2003, 13:40:16  
Job time : 1621.04 secs

**THIS PAGE BLANK (USPTO)**



score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 12:10:39 ; Search time 4099.22 Seconds (without 31:44:47)

9378.574 million cell updates/sec

Title: US-09-934-249-1

Sequence: 1 cgaccgcgtctcgcgagcga.....ctgcgtaggtgaagaagcag 1321

Scoring table: IDENTITY\_NUC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 08

Listing first 45 summaries

Database :

1:	GenEmb1: *
2:	gb_ba: *
3:	gb_hig: *
4:	gb_in: *
5:	gb_om: *
6:	gb_ov: *
7:	gb_pat: *
8:	gb_ph: *
9:	gb_pl: *
10:	gb_pr: *
11:	gb_rn: *
12:	gb_scs: *
13:	gb_sy: *
14:	gb_un: *
15:	gb_vl: *
16:	em_ba: *
17:	em_fun: *
18:	em_hum: *
19:	em_in: *
20:	em_om: *
21:	em_ov: *
22:	em_pat: *
23:	em_ph: *
24:	em_pl: *
25:	em_rn: *
26:	em_scs: *
27:	em_sy: *
28:	em_un: *
29:	em_vl: *
30:	em_hig_hum: *
31:	em_hig_in: *
32:	em_hig_other: *
33:	em_hig_om: *
34:	em_hig_pat: *
35:	em_hig_ph: *
36:	em_hig_pl: *
37:	em_hig_rn: *
38:	em_hig_sy: *
39:	em_hig_un: *
40:	em_hig_vl: *
41:	em_hig_hum: *

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	1321	100.0	1321	6	AX392417	AX392417 Sequence
2	1229	93.0	4839	6	AF305616	AF305616 Homo sapi
3	887.4	67.2	1061	6	BC015918	BC015918 Homo sapi
4	861	65.2	861	6	AX392419	AX392419 Sequence
5	802.2	60.7	1141	9	AF224278	AF224278 Homo sapi
6	800	60.6	1818	6	AY128643	AY128643 Homo sapi
7	616.4	46.7	878	6	AX392428	AX392428 Sequence
8	594.4	45.0	61505	9	AF305426	AF305426 Homo sapi
9	594.4	45.0	130435	9	AL103541	AL103541 Human DNA
10	532.8	37.4	150224	9	HSJ1059L7	HSJ1059L7 Human DNA
11	493.6	39.6	693	6	AX392430	AX392430 Sequence
12	425	32.2	651	10	AF220208	AF220208 Mus muscu
13	417.8	31.6	812	6	AX011709	AX011709 Sequence
14	401.4	30.4	408	6	AX071267	AX071267 Sequence
15	362	27.4	156075	2	AL837509	AL837509 Mus muscu
16	352.2	26.7	8494	9	AF009424	AF009424 Homo sapi
17	350	26.5	2170	9	AK055028	AK055028 Homo sapi
18	349	26.4	8093	6	AX392432	AX392432 Sequence
19	349	26.4	8093	6	AF009426	AF009426 Homo sapi
20	341.8	25.9	475	6	AX392431	AX392431 Sequence
21	284.2	21.5	8440	6	AF009425	AF009425 Homo sapi
22	281	21.3	8039	9	AF009427	AF009427 Homo sapi
23	250.2	18.9	172692	9	AF001013	AF001013 Homo sapi
24	250.2	18.9	173709	9	AP001010	AP001010 Homo sapi
25	250.2	18.9	181083	2	AP001268	AP001268 Homo sapi
26	240.8	18.2	766	10	BC022716	BC022716 Mus muscu
27	232.6	17.6	66972	2	AC1117364	AC1117364 Rattus no
28	231	17.5	183619	2	AC1111069	AC1111069 Mus muscu
29	228.2	17.3	150542	2	AC097603	AC097603 Rattus no
30	215.8	16.3	167758	2	AC110189	AC110189 Mus muscu
31	156.2	11.8	167758	2	AC110189	AC110189 Mus muscu
32	120	9.1	159824	2	AC111878	AC111878 Rattus no
33	102	7.7	161773	2	AC1103491	AC1103491 Rattus no
34	101.6	7.7	130626	2	AC113687	AC113687 Rattus no
35	99.8	7.6	207420	2	AC078884	Mus muscu
36	99.6	7.5	108782	2	AC123508	Rattus no
37	99.4	7.5	146091	2	AC099260	Rattus no
38	99	7.5	83629	2	AC110485	Rattus no
39	98.8	7.5	161773	2	AC103491	Rattus no
40	96.4	7.3	154684	2	AC128630	Rattus no
41	96	7.3	190457	2	AC119326	Rattus no
42	94.4	7.1	208757	2	AC094591	Rattus no
43	94.2	7.1	147622	2	AC110437	Rattus no
44	94	7.1	151283	2	AC107540	Rattus no
45	93.8	7.1	136014	2	AC126531	Rattus no

JOURNAL Patent: WO 0216416-A 1 28-FEB-2002;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)  
FEATURES Location/Qualifiers  
SOURCE 1. 1321  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

## CDS

413. 1276  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD29005.1"  
/db\_xref="GI:19700733"  
/translation="MHRLMGYNSTAAAGAPNVCNCKRSIFQSMETLEFVQI  
IIVVMVMVAVITCLSHYKLSRSTFISHSGRRREDALSEGLMSESVGN  
GIPPOYAPRPDPLAAPPFAQERFHRQPTPYLQHDIDLPITISLDEPEPP  
YQGPCTIOLRPEQOLEINRESVAPRNRIFFSDLMDSALHGGPCPPSSNSISATP  
YSGRMGPPPTSEVYGHYPGSSFOHOSSPPSLIEGRLHHTHIAPLESAIWS  
KEKDKOKHPL"

BASE COUNT 223 a 493 c 440 g 165 t  
ORIGIN  
Query Match 100.0%; Score 1321; DB 6; Length 1321;  
Best Local Similarity 100.0%; Pred. No. 9.5e-176;  
Matches 1321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACGCGGTCTCGAGCAAAACCGATCTCTTGACTTGATAGAGAGAGAGGGG 60  
DB 1 CGACGCGGTCTCGAGCAAAACCGATCTCTTGACTTGATAGAGAGAGAGGGG 60  
QY 61 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
DB 61 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
QY 121 GCGGCGAGCG 180  
DB 121 GCGGCGAGCG 180  
QY 121 GCGGCGAGCG 180  
DB 121 GCGGCGAGCG 180  
QY 181 AGCG 240  
DB 181 AGCG 240  
QY 241 TGAGCG 300  
DB 241 TGAGCG 300  
QY 301 CTCCCG 360  
DB 301 CTCCCG 360  
QY 361 CCG 420  
DB 361 CCG 420  
QY 421 CTGATGAGGGGTCAACAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
DB 421 CTGATGAGGGGTCAACAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
QY 481 GTGCAAGTGAACAGCTCTTTGTTCCAGAGCATGAGATACGAGAGTGGAGTTTCA 540  
DB 481 GTGCAAGTGAACAGCTCTTTGTTCCAGAGCATGAGATACGAGAGTGGAGTTTCA 540  
QY 541 GATCATCATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 541 GATCATCATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
QY 601 CCACTACAGTGTCTGACAGGTCTCTTATACAGCGGCAACAGAGGGGAGAGAGA 660  
DB 601 CCACTACAGTGTCTGACAGGTCTCTTATACAGCGGCAACAGAGGGGAGAGAGA 660  
QY 661 AGATGCGCTGTCTCAGAGAGATGCTGTGCGCTCGAGAGCAAGTGTGAGCAACGG 720  
DB 661 AGATGCGCTGTCTCAGAGAGATGCTGTGCGCTCGAGAGCAAGTGTGAGCAACGG 720  
QY 721 AATCCAGAGCGCGAGTCTAGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780

DB 721 AATCCAGAGCGCGAGTCTAGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 780  
QY 781 CTGCGCCAGCGGAGAGCGCTTCCAGCGCTTCCAGCGCCACTATCCGTACTGAGCAGA 840  
DB 781 CTGCGCCAGCGGAGAGCGCTTCCAGCGCTTCCAGCGCCACTATCCGTACTGAGCAGA 840  
QY 841 GATGACCTGCGACCCCACTATCTGCTGATGACAGGGGAGAGAGAGAGAGAGAG 900  
DB 841 GATGACCTGCGACCCCACTATCTGCTGATGACAGGGGAGAGAGAGAGAGAGAG 900  
QY 901 CCCCTGACCCCTCCAGCTTGGGAGCCCGAGACAGCTGGAACCTAACCAGGAGTCGT 960  
DB 901 CCCCTGACCCCTCCAGCTTGGGAGCCCGAGACAGCTGGAACCTAACCAGGAGTCGT 960  
QY 961 GCGCGCACCCCAACAGAACCACTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020  
DB 961 GCGCGCACCCCAACAGAACCACTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020  
QY 1021 CGGCG 1080  
DB 1021 CGGCG 1080  
QY 1081 GCGCATGAGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140  
DB 1081 GCGCATGAGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140  
QY 1141 CTTCAGACACAG 1200  
DB 1141 CTTCAGACACAG 1200  
QY 1201 CACACACATCCG 1260  
DB 1201 CACACACATCCG 1260  
QY 1261 AGGACACCTCTCTTAGGTTCCCGAGGGGGCGCGCGCGCGCGCGCGCGCGCG 1320  
DB 1261 AGGACACCTCTCTTAGGTTCCCGAGGGGGCGCGCGCGCGCGCGCGCGCGCG 1320  
QY 1321 G 1321  
DB 1321 G 1321

RESULT 2  
AF305616 4839 bp mRNA linear PRI 21-OCT-2001  
LOCUS Homo sapiens STAG1/PMEP1 mRNA, complete cds.  
ACCESSION AF305616  
VERSION AF305616.1 GI:16303741  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 (bases 1 to 4839)  
Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.  
Characterization of a novel gene, STAG1/PMEP1, upregulated in  
renal cell carcinoma and other solid tumors  
Mol. Carcinog. 32 (1), 44-53 (2001)  
JOURNAL  
MEDLINE  
PUBMED  
2 (bases 1 to 4839)  
Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.  
Direct Submision  
Submitted (14-SEP-2000) Centre for Molecular Biotechnology, QLD  
Queensland University of Technology, 2 George St, Brisbane, QLD  
4001, Australia

FEATURES  
SOURCE 1. 4839  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="20"

Query Match	93.0%;	Score 1229;	DB 9;	Length 4839;
Best Local Similarity	100.0%;	Pred. No. 5.6e-163;		
Matches 1229;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Db	721	AGGACCTATTCGTA	CTGCAGCAGCAG	ATGCACTTCCAC	CCACCATCTTCG	CTGTCAG	780
OY	873	ACGGGAGAGACCCC	CCACCTTACCA	AGGGCCCCCTG	CACCTTCAGCTT	CGSACCCGAGC	932
Db	781	ACGGGAGAGACCCC	CCACCTTACCA	AGGGCCCCCTG	CACCTTCAGCTT	CGSACCCGAGC	840
OY	933	AGCAGCTGAACTGA	CAACCGGGAGT	CGGCGGACCC	CCAAACGAA	CACTCTTCGACA	992
Db	841	AGCAGCTGAACTGA	CAACCGGGAGT	CGGCGGACCC	CCAAACGAA	CACTCTTCGACA	900
OY	993	GTCACCTGATGAT	AGTCCACAGG	CTGGGCCCCCT	GCCTCCCGCAG	CACTAACTGGGCA	1052
Db	901	GTCACCTGATGAT	AGTCCACAGG	CTGGGCCCCCT	GCCTCCCGCAG	CACTAACTGGGCA	960
OY	1053	TCAGCGCCACGTC	GTACGGCAGCG	CGCGGGCGCAT	GGAGGGGCGCG	CGCCGCGCATTCACAG	1112
Db	961	TCAGCGCCACGTC	GTACGGCAGCG	CGCGGGCGCAT	GGAGGGGCGCG	CGCCGCGCATTCACAG	1020
OY	1113	AGGTCATCGGCG	CACCTACCGGGG	GTCTCTTCGAC	AGCACCAGCAG	CAAGCAGTGGCGCGCCT	1172
Db	1021	AGGTCATCGGCG	CACCTACCGGGG	GTCTCTTCGAC	AGCACCAGCAG	CAAGTGGCGCGCCT	1080
OY	1173	CCTTCCTGAGAGG	AGGAGCCGGG	CTCCACACAC	ACATGGCGCCCT	TAAGAGCGCAGCA	1232
Db	1081	CCTTCCTGAGAGG	AGGAGCCGGG	CTCCACACAC	ACATGGCGCCCT	TAAGAGCGCAGCA	1140
OY	1233	TCTGGAGCAAA	GAGAGATTAAC	AGAAAGACA	CCCTCTAGGGT	TCGCCAGGGGGCC	1292
Db	1141	TCTGGAGCAAA	GAGAGATTAAC	AGAAAGACA	CCCTCTAGGGT	TCGCCAGGGGGCC	1200
OY	1293	GGGCTGGGGG	CTGCTAGT	AGTGAAGG	CAG	1321	
Db	1201	GGGCTGGGGG	CTGCTAGT	AGTGAAGG	CAG	1229	

[illegible]

Location/Qualifiers

1. 1061

CD5

ORGANISM

REFERENCE  
AUTHORS

TITLE

## FEATURES

CDS

BASE COUNT	172 a	307 c	229 g
ORIGIN			

Query Match	65.46%	Score 607	
Best Local Similarity	100.08%	Pred. No. 2.3e-111	
Matches	861	Mismatches	0
		Indels	0
		Gaps	0

QY 41

QY 47

QY 53

500Y

Db 18

	gene	/db_xref=taxon:9606"
		/chromosome=20"
		/map=20q13.31-13.33"
		/cell_line=LNCap"
		1..1141
CDS		/gene="PMEP1"
		96..854
		/note="Type Ib transmembrane protein; expression is induced in response to the synthetic androgen, R1881; expression is abundant in, and restricted to, prostate glandular epithelial cells; similar to the predicted protein encoded by sequence deposited at GenBank accession Number NP_004329"
		/codon_start=1
		/product="PMEP1 protein"
		/protein_id="AAF6322.1"
		/db_xref="GI:9255809"
		/translation="MAELEFVQIIIVVMVVYITCLLSHYKLSARSFISHSOGRRRLRDLSECCIMPSEIVSNGIIPROYAVRPDRLAVPPAFORERHROPPTYLYOHEIDLPPTILSDGEERPYPQCOTLLRPEDOLELRVSVAAPNRTFDSDLMDRAIGDPCPPSSNSGISATCYGSGSRMGEGPPYSEVIYHPGSSFQHQQSGPRPSLEGTIRHHHTNPILPSAATWSTEXKKOKHPL"
BASE COUNT	271 a	350 c
ORIGIN		336 g
		184 t
Query Match	60.7%;	Score 802.2;
Best Local Similarity	98.4%;	Pred. No. 3.7e-103;
Matches	810;	Conservative 0;
	Mismatches	13;
	Indels	0;
	Gaps	0;
Qy	499 TTTGTTCCAGAGCATGAGATCAGACCTGGAGCTTGTTCAGATCATCATCGTG	558
Db	77 TCCTCCGGAACCAGCAATGCAGACTGGAGTTGTTCAGATCATCATCGTG	136
Qy	555 GGTGATGATGCGTAGTGCGTGCGTATCATACGTCGCTCTGAGGCCATCAAGCTGTG	618
Db	137 GGTTGATGATGGTGAAGTGCGTGCGTATCATACGTCGCTCTGAGGCCATCAAGCTGTG	196
Qy	619 ACGGTCCTTCATCACCCGGCAGCCAGCCAGGGGGGAGAGAGAAGATGCTGTCCTGACA	678
Db	197 ACGGTCCTTCATCACCCGGCAGCCAGCCAGGGGGGAGAGAGAAGATGCTGTCCTGACA	256
Qy	679 AGGATGCTGTGGCCCTCGGAGAGCAGATGTCAGGCAAGGAAATCCAGAGCCGACAGT	738
Db	257 AGGATGCTGTGGCCCTCGGAGAGCAGATGTCAGGCAAGGAAATCCAGAGCCGACAGT	316
Qy	739 CTACGCCCGCGCTCGGGCCACCGACCGCTGGCCCTGCGCCCTTCGCGCCAGCGGAGCG	798
Db	317 CTACGCCCGCGCTCGGGCCACCGACCGCTGGCCCTGCGCCCTTCGCGCCAGCGGAGCG	376
Qy	799 CTTGCACGCGCTTCAGGCCACCTATTCGTAACCTGAGACAGGATCGACCTGCGCAC	858
Db	377 CTTGCACGCGCTTCAGGCCACCTATTCGTAACCTGAGACAGGATCGACCTGCGCAC	436
Qy	859 CATCTCGCTGTGACAGGGGAGAGACCCCCACACCTACAGAGGCCCCCTGACCTTCAGCT	918
Db	437 CATCTCGCTGTGACAGGGGAGAGACCCCCACACCTACAGAGGCCCCCTGACCTTCAGCT	496
Qy	919 TTGGGACCCCGAGCAGCAGCTGGAATCAACCGGAGTGTGGCGCGCACCCCAAACAG	978
Db	497 TTGGGACCCCGAGCAGCAGCTGGAATCAACCGGAGTGTGGCGCGCACCCCAAACAG	556
Qy	979 AACCATCTTCGACAGTGAACCTGATGATAGTCCAGAGGCTGGGCGGCCCCCCCAG	1038
Db	557 AACCATCTTCGACAGTGAACCTGATGATAGTCCAGAGGCTGGGCGGCCCCCCCAG	616
Qy	1039 CAGTAACCTCGGACATCAGCGCACCTGTCTACGGCAGCGCGGCGCATGAGAGGGCCGCC	1098
Db	617 CAGTAACCTCGGACATCAGCGCACCTGTCTACGGCAGCGCGGCGCATGAGAGGGCCGCC	676
Qy	1099 GCCCACCTACAGCGAGGTGATCGGCCACTACCCGGGGCTCTCTTCACGACACAGACAG	1158
Db	677 GCCCACCTACAGCGAGGTGATCGGCCACTACCCGGGGCTCTCTTCACGACACAGACAG	736

QY 1159 CAGTGGGCGCCCTCTGCTGAGAGGACCGGCTCCACCAACACATCGGCCCT 1218  
 DB 737 CAGTGGGCGCCCTCTGCTGAGAGGACCGGCTCCACCAACACATCGGCCCT 796  
 QY 1219 AGAGAGCCGACCATCTGAGCAAGAGAGATTAACAGAAAGAGACCTCTCTAGG 1278  
 DB 797 AGAGAGCCGACCATCTGAGCAAGAGAGATTAACAGAAAGAGACCTCTCTAGG 856  
 QY 1279 TCCCGAGGGGCGGCGCTGGGCGCTGCTAGTGAAGAGAG 1321  
 DB 857 TCCCGAGGGGCGGCGCTGGGCGCTGCTAGTGAAGAGAG 899

# RESULT 6

LOCUS AY128643 1818 bp mRNA linear PRI 05-AUG-2002  
 DEFINITION Homo sapiens PMP1 variant A protein mRNA, complete cds.  
 ACCESSION AY128643  
 VERSION 1  
 KEYWORDS AY128643.1 GI:22121998

## ORGANISM

human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1818)  
 Brunschwig, E.B., Wilson, K., Mack, D., Dawson, D., Lawrence, E.,  
 Willson, J.K.V., Lu, S., Nostri, A., Swinler, S., Beard, L.,  
 Lutterbaugh, J.D., Willis, J., Platzer, P. and Markowitz, S.  
 Submitted (03-JUL-2002) Department of Medicine, Case Western  
 Reserve University/Howard Hughes Medical Institute, 11001 Cedar  
 Ave., Cleveland, OH 44106, USA

## REFERENCE

TITLE JOURNAL  
 Location/Qualifiers  
 1. 1818  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20q13.31-13.33"  
 145..858  
 /codon\_start=1  
 /product="PMP1 variant A protein"  
 /protein\_id="AA089277.1"  
 /db\_xref="GI:22121998"  
 /translation="MMWVYVITCLISYKISARSFISRSQGRREDALISSEGLMP  
 SESTYSGNPEPVYVAPPRPTDLAPPPRQRRFRFPDTPYVYLQHEIDLPTTSL  
 SDGEPPYQGPCTIQLEDPQLELNRESVAPRNRTIFDSIDMSARLGGPCPPSS  
 NSGISATCTGSGGNGEPPPYSEVIGHYPSFSFOHOSSGPPSLBGLRLHHTHLP  
 LESAIWSEKDKQKHPL"

## FEATURES

SE COUNT 459 a 467 c 494 g 398 t  
 ORIGIN

Query Match 60.6%; Score 800; DB 9; Length 1818;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-103;  
 Matches 800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 522 CGAGCTGAGTTTGTTCAGATCATCATCTGCTGATGATGATGATGATGATG 581  
 DB 104 CGAGCTGAGTTTGTTCAGATCATCATCTGCTGATGATGATGATGATGATG 163  
 QY 582 TGATACGTCGCTGTGAGCACTAAGCTGTCTGCACGCTCTTCATCAGCGGACA 641  
 DB 164 TGATACGTCGCTGTGAGCACTAAGCTGTCTGCACGCTCTTCATCAGCGGACA 223  
 QY 642 GCCAGGGGCGGAGAGAGATGCGCTGTCTCAGAAAGATGCTGCGGCGCTCGGAGA 701  
 DB 224 GCCAGGGGCGGAGAGAGATGCGCTGTCTCAGAAAGATGCTGCGGCGCTCGGAGA 283  
 QY 702 GCACAGTGTGAGGAGAGATGCCAGAGCGGAGTCTACGCCCGCTCGGCGCACCG 761  
 DB 284 GCACAGTGTGAGGAGAGATGCCAGAGCGGAGTCTACGCCCGCTCGGCGCACCG 343  
 QY 762 ACCGCTGCGCGCTCGCGCGCTCGCGAGGAGCGCTTCACCGCTTCACAGCCACT 821

DB 344 ACCGCTGCGCGCTCGCGCGCTTCGCGAGCGGAGGGCTTCACCGCTTCACAGCCACT 403  
 QY 822 ATCCGTAACCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 881  
 DB 404 ATCCGTAACCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 463  
 QY 882 AGCCCCACCTTACAGAGGCGCCCTGACCTTCAGCTTGGGAGCCCGAGCAGCTGG 941  
 DB 464 AGCCCCACCTTACAGAGGCGCCCTGACCTTCAGCTTGGGAGCCCGAGCAGCTGG 523  
 QY 942 AACTGAAACCGGAGTGTGGTGGGCGCAGCCCAACAGAACATCTTGCAGTACCTGA 1001  
 DB 524 AACTGAAACCGGAGTGTGGTGGGCGCAGCCCAACAGAACATCTTGCAGTACCTGA 583  
 QY 1002 TGGATAGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1061  
 DB 584 TGGATAGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 643  
 QY 1062 CGTGTACGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1121  
 DB 644 CGTGTACGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 703  
 QY 1122 GCCACTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1181  
 DB 704 GCCACTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 763  
 QY 1182 AGGGAGCCGCGCTCCACACACACATCGGCGGCGGCGGCGGCGGCGGCGGCGG 1241  
 DB 764 AGGGAGCCGCGCTCCACACACACATCGGCGGCGGCGGCGGCGGCGGCGGCGG 823  
 QY 1242 AAGAGAGGATTAACAGAAAGAGACCTCTCTAGGCTCCCGAGGCGGCGGCGGCGG 1301  
 DB 824 AAGAGAGGATTAACAGAAAGAGACCTCTCTAGGCTCCCGAGGCGGCGGCGGCGG 883  
 QY 1302 CTGCGTAGGTGAAGAGCAG 1321  
 DB 884 CTGCGTAGGTGAAGAGCAG 903

## RESULT 7

LOCUS AX392428 878 bp DNA linear PAT 23-MAR-2002  
 DEFINITION Sequence 12 from Patent WO0216416.  
 ACCESSION AX392428  
 VERSION 1  
 KEYWORDS AX392428.1 GI:19700744

## SOURCE

house mouse.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murine; Mus.  
 1  
 Lee, R.T., Landschulz, K.T., Kennedy, S.P., Thompson, J.F. and  
 Turi, T.G.  
 Diagnosis and treatment of cardiovascular conditions  
 Patent: WO 0216416-A 12 28-FEB-2002;  
 THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)

## FEATURES

source  
 1..878  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 20..844  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD29011.1"  
 /db\_xref="GI:19700745"  
 /translation="MGVGTAAAGAPNVCACNCRSIFPSMETTELEFVIVIV  
 VVAVVWVITCLISYKISARSFISRSQGRREDALISSEGLMPSESTYSGNPEPVYVAPPRPTDLAPPPRQRRFRFPDTPYVYLQHEIDLPTTSL  
 SDGEPPYQGPCTIQLEDPQLELNRESVAPRNRTIFDSIDMSARLGGPCPPSSNSGISATCTGSGGNGEPPPYSEVIGHYPSFSFOHOSSGPPSLBGLRLHHTHLP  
 LESAIWSEKDKQKHPL"

## CDS

BASE COUNT 179 a 305 c 247 g 147 t  
 ORIGIN

[illegible]

KEYWORDS	source	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	source
Homo sapiens.		Homo sapiens										
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.												
1 (bases 1 to 61505)												
Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.												
Identification and characterization of a novel gene, STAGL, up-regulated in renal cell carcinoma and other solid tumours												
Unpublished												
2 (bases 1 to 61505)												
Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.												
Direct Submission												
Submitted (13-sep-2000) Centre for Molecular Biotechnology, Queensland University of Technology, 2 George St, Brisbane, QLD 4001, Australia												
Location/Qualifiers												
1..61505												
/organism="Homo sapiens"												
/db_xref="taxon:9606"												
/chromosome="20"												
/map="20q13.2-q13.33"												
1..61505												
/gene="STAG1/PMEPA1"												
/join(1..429,50206..50360,56817..56870,57305..61505)												
/gene="STAG1/PMEPA1"												
/product="solid tumor-associated 1 protein"												
join(331..429,50206..50360,56817..56870,57305..57850)												
/gene="STAG1/PMEPA1"												
/codon_start=1												
/product="solid tumor-associated 1 protein"												
/protein_id="AA09357.1"												
/db_xref="GI:15824469"												
/translation="MHRLMGVNSTAAAGAPVNSCTCNCKRSLFOSMETLEEFVOII IIVVMMVVVITCLISHLKLSASPSISHSGRRREDAISEGLMSESVSGN GIPEPVYAPRPPTDRLIAYPEFAERERHNPQPIYPLQIHEDIPRTISLSDGEPP YQGCTQLRDPEQDLELNRSVRAIPKRTIFDSLDMSARLGPCPPSSSGISATC YGSGRMREGPPPTSEVYIGHVPGSSIFQHODSSGPPSILEGRLHHTHIAPLESAIWS KENDKRGHP"												
BASE COUNT 13329 a 16074 c 17430 g 14672 t												
ORIGIN												
Query Match 45.0%; Score 594.4; DB 9; Length 61505;												
Best Local Similarity 99.0%; Pred. No. 2.2e-74;												
Matches 598; Conservative 0; Mismatches 6; Indels 0; Gaps 0;												
718 CGGATCCAGAGCGCGAGTCTACGCCCGCTGGCCACGACGCGTGGCGTGC												
57292 CTGCTTCTCCAGCCGCGAGTCTACGCCCGCTGGCCACGACGCGTGGCGTGC												
778 GCCCTTCCGCGCGGAGCGCTTCCACGCTTCCAGCCACCTATCCGCTACGACA												
57352 GCCCTTCCGCGCGGAGCGCTTCCACGCTTCCAGCCACCTATCCGCTACGACA												
838 CGAGATTCAGACCTCGGACCCACCATCTGCTGTCACAGCGGAGGAGGCCCCACCTACCA												
57412 CGAGATTCAGACCTCGGACCCACCATCTGCTGTCACAGCGGAGGAGGCCCCACCTACCA												
898 GGGCCCTTGCACACCTTCAGCTTGGGACCCCGAGAGCAGCTGGAACTGAACCGGAGATC												
57472 GGGCCCTTGCACACCTTCAGCTTGGGACCCCGAGAGCAGCTGGAACTGAACCGGAGATC												
958 GGTGGCGGCGACCCCAACAGACCATCTTGCACAGTGAAGCTGATGATGTCAGAGCT				</								

QY 1138 CTCCTTCAGACACAGAGAGAGTGGGGCCCTCTTCTGAGAGGAGCCCGCTCCA 1197  
 |||||  
 Db 57712 CTCCTTCAGACACAGAGAGAGTGGGGCCCTCTTCTGAGAGGAGCCCGCTCCA 57771  
 QY 1198 CCACACACACATCGGCCCTAGAGAGCGACCATCTGAGCAAGAAGAGGATTAACA 1257  
 |||||  
 Db 57772 CCACACACACATCGGCCCTAGAGAGCGACCATCTGAGCAAGAAGAGGATTAACA 57831  
 QY 1258 GAAGAGACACCTCTCTAGAGGTCCCGAGGGGGCCGCGCTGGTAGTGAAAAG 1317  
 |||||  
 Db 57832 GAAGAGACACCTCTCTAGAGGTCCCGAGGGGGCCGCGCTGGTAGTGAAAAG 57891  
 QY 1318 GCAG 1321  
 |||||  
 Db 57892 GCAG 57895

RESULT 9  
 HS718J7/c  
 DEFINITION Human DNA sequence from clone RP4-718J7 on chromosome 20q13.31-13.33 contains the PKC1 gene for soluble phosphoenolpyruvate carboxykinase 1, part of a novel gene similar to mouse DLM-1 (tumour stroma and activated macrophage protein), the 3' end of the TMEPAI gene encoding an androgen induced 1b transmembrane protein (PMEPAI), two putative novel genes, a Cpg island, ESTs, STS and GSSs, complete sequence.  
 AL035541  
 AL035541.15 GI:11546043  
 HG: CPG island; DLM-1; macrophage protein; PKC1; phosphoenolpyruvate carboxykinase; PMEPAI; transmembrane protein.

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 130435)  
 AUTHOR Sehra, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, UK  
 REFERENCE Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, UK  
 COMMENT Request: clonerequest@sanger.ac.uk  
 On Dec 5, 2000 this sequence version replaced gi:10198628.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission only a small overlap as described above, as we submit sequences with numbers given in the feature table with their source databases:  
 Em: EMBL, Sw: SWISSPROT, Tr: TrEMBL, Wp: WORMPEP, Information on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr20

This sequence is the entire insert of clone RP4-718J7 the true left right end of clone RP5-1007E6 is at 71437 in this sequence. The true right end of clone RP4-579P20 is at 43945 in this sequence. RP4-718J7 is from the library RPCI-4 constructed by the group of Plietier de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm  
 VECTOR: pCYPAC2

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES  
 source  
 1..130435  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="q13.31-13.33"  
 /clone="RP4-718J7"  
 /clone\_lib="RPCI-4"  
 319..478  
 /note="5 copies 32 mer 67% conserved"  
 370..463  
 /note="2 copies 47 mer 87% conserved"  
 427..626  
 /note="4 copies 50 mer 71% conserved"  
 707..812  
 /note="53 copies 2 mer cc 61% conserved"  
 967..2752  
 /note="893 copies 2 mer gg 54% conserved"  
 982..2730  
 /note="33 copies 53 mer 54% conserved"  
 1177..1820  
 /note="4 copies 161 mer 64% conserved"  
 1204..1371  
 /note="3 copies 56 mer 75% conserved"  
 1358..2757  
 /note="28 copies 50 mer 54% conserved"  
 1675..2718  
 /note="18 copies 58 mer 55% conserved"  
 1843..2345  
 /note="9 copies 56 mer 64% conserved"  
 1852..2079  
 /note="4 copies 57 mer 86% conserved"  
 2028..2632  
 /note="11 copies 55 mer 60% conserved"  
 2087..2257  
 /note="3 copies 57 mer 83% conserved"  
 2228..2497  
 /note="5 copies 54 mer 75% conserved"  
 2583..2716  
 /note="2 copies 67 mer 82% conserved"  
 3378..3426  
 /note="LIMB4 repeat: matches 6088..6136 of consensus"  
 3799..3896  
 /note="Charliel repeat: matches 681..781 of consensus"  
 5331..5793  
 /note="WTR1B repeat: matches 14..466 of consensus"  
 5797..5988  
 /note="6 copies 32 mer 86% conserved"  
 5855..5978  
 /note="31 copies 4 mer gcac 61% conserved"  
 6474..6591  
 /note="MFR repeat: matches 91..218 of consensus"  
 6592..6723  
 /note="FLAM C repeat: matches 1..132 of consensus"  
 6952..7021  
 /note="L2 repeat: matches 2637..2705 of consensus"  
 7358..7671  
 /note="AluXs repeat: matches 1..312 of consensus"  
 8521..8554  
 /note="Alu repeat: matches 1..34 of consensus"  
 8863..9217  
 /note="MUTID repeat: matches 105..505 of consensus"  
 9228..9746  
 /note="L2 repeat: matches 1459..2028 of consensus"  
 9880..10089  
 /note="L2 repeat: matches 2292..2511 of consensus"  
 10094..10206  
 /note="AluDb repeat: matches 188..300 of consensus"  
 10263..10573  
 /note="AluDb repeat: matches 1..308 of consensus"  
 10658..10753  
 /note="L2 repeat: matches 2647..2749 of consensus"  
 11816..12380



```

repeat_region      /note="LTR19B repeat: matches 1. .580 of consensus"
12519. .12813
repeat_region      /note="L1MD2 repeat: matches 6032. .6331 of consensus"
13559. .13726
repeat_region      /note="3 copies 56 mer 73% conserved"
13585. .13716
repeat_region      /note="22 copies 6 mer ctctct 67% conserved"
13586. .13717
repeat_region      /note="66 copies 2 mer tc 68% conserved"
13588. .13715
repeat_region      /note="4 copies 32 mer 78% conserved"
13589. .13724
repeat_region      /note="34 copies 4 mer cttt 77% conserved"
complement(13622. .14142)
misc_feature        /note="match: GSS: Em:AQ592603"
14313. .14437
repeat_region      /note="LTR16C repeat: matches 257. .386 of consensus"
15015. .15110
repeat_region      /note="3 copies 32 mer 79% conserved"
15021. .15110
repeat_region      /note="45 copies 2 mer ac 80% conserved"
15026. .15109
repeat_region      /note="7 copies 12 mer 82% conserved"
15028. .15111
repeat_region      /note="14 copies 6 mer cacaca 82% conserved"
15029. .15108
repeat_region      /note="20 copies 4 mer acac 83% conserved"
15273. .15399
repeat_region      /note="L2 repeat: matches 2159. .2285 of consensus"
15420. .15721
repeat_region      /note="AluXx repeat: matches 1. .303 of consensus"
16333. .16396
repeat_region      /note="2 copies 32 mer 98% conserved"
16573. .16690
repeat_region      /note="L2 repeat: matches 2596. .2711 of consensus"
18049. .18169
repeat_region      /note="MIR repeat: matches 86. .211 of consensus"
18312. .18438
repeat_region      /note="MIR repeat: matches 138. .250 of consensus"
18585. .18776
repeat_region      /note="MER20 repeat: matches 7. .217 of consensus"
19211. .19290
repeat_region      /note="MIR repeat: matches 72. .154 of consensus"
20413. .20463
repeat_region      /note="MIR repeat: matches 90. .143 of consensus"
complement(21005. .21672)
misc_feature        /note="match: GSS: Em:AQ748384"
complement(21187. .21672)
misc_feature        /note="match: GSS: Em:AQ776209"
complement(21278. .21663)
misc_feature        /note="match: GSS: Em:AQ136459"
complement(21392. .21672)
misc_feature        /note="match: GSS: Em:AQ892114"
21667. .21775
repeat_region      /note="MIR repeat: matches 73. .192 of consensus"
22577. .23646
repeat_region      /note="2 copies 35 mer 100% conserved"
24257. .24591
repeat_region      /note="L2 repeat: matches 2336. .2710 of consensus"
25580. .25611
repeat_region      /note="16 copies 2 mer tc 87% conserved"
26333. .26643
repeat_region      /note="AluY repeat: matches 1. .310 of consensus"
27603. .27684
repeat_region      /note="L2 repeat: matches 2661. .2739 of consensus"
27823. .28041
repeat_region      /note="MIR repeat: matches 38. .242 of consensus"
28129. .28259
repeat_region      /note="F1AM_C repeat: matches 1. .127 of consensus"
31203. .31356
repeat_region      /note="L1MB5 repeat: matches 6015. .6176 of consensus"
32224. .32522
repeat_region      /note="AluXx repeat: matches 1. .298 of consensus"

```

```

repeat_region      33285. .33481

Query Match      45.0%; Score 594.4; DB 9; Length 130435;
Best Local Similarity 99.0%; Pred. No.1.9e-74;
Matches 598; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 718 CGGAATCCAGAGCCGAGGCTCAGGCGCCCGCCGACCGACGACCGCTGGCCGTGCC 777
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128454 CTGCTTCCTCCAGCCGAGAGGTCTACGCCGCCGCTGGCCACGACCGCTGGCCGTGCC 128395

QY 778 GCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCGACCTATCCGTAAGCTGCAGCA 837
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128394 GCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCGACCTATCCGTAAGCTGCAGCA 128335

QY 838 CGAGATGACCTGCGACCCACCATCTGCTGTACAGAGGGAGAGGCCGCCACCTTACCA 897
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128334 CGAGATGACCTGCGACCCACCATCTGCTGTACAGAGGGAGAGGCCGCCACCTTACCA 128275

QY 898 GGGGCCCCGACCCCTCAGGTTGGGAGCCCGACGACGACGCTGGAACCGGGAGTC 957
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128274 GGGGCCCCGACCCCTCAGGTTGGGAGCCCGACGACGACGCTGGAACCGGGAGTC 128215

QY 958 GGTGCGGCAACCCCAACAGAACCATCTTCGACAGTGACCTGATGATGTCAGAGCT 1017
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128214 GGTGCGGCAACCCCAACAGAACCATCTTCGACAGTGACCTGATGATGTCAGAGCT 128155

QY 1018 GGGGCGCCCTGCCCCCGCCAGCACTGAGGCGATAGCGCCACGTCGTAAGCGACCG 1077
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128154 GGGGCGCCCTGCCCCCGCCAGCACTGAGGCGATAGCGCCACGTCGTAAGCGACCG 128095

QY 1078 CGGGCGCATGAGGGGGGGGGGGCGCCGACCTACAGAGGATGATGGCGACACCGGGGTC 1137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128094 CGGGCGCATGAGGGGGGGGGGGCGCCGACCTACAGAGGATGATGGCGACACCGGGGTC 128035

QY 1138 CTCTTCACGACACGACGAGAGCACTGGCGCCCTCTCTGCTGAGGGAGCCCGCTCA 1197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128034 CTCTTCACGACACGACGAGAGCACTGGCGCCCTCTCTGCTGAGGGAGCCCGCTCA 127975

QY 1198 CCACACACATCGCGCCCTAGAGAGCGACCATCTGGAGCAAGAAGATTAACA 1257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127974 CCACACACATCGCGCCCTAGAGAGCGACCATCTGGAGCAAGAAGATTAACA 127915

QY 1258 GAAAGGACACCCCTCTAGGGTCCCAAGGGGGGGCGGGCTGCGGCTGTAAGGAAAG 1317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127914 GAAAGGACACCCCTCTAGGGTCCCAAGGGGGGGCGGGCTGCGGCTGTAAGGAAAG 127855

QY 1318 GCAG 1321
    |||||
Db 127854 GCAG 127851

RESULT 10
HSJ1059L7/c      150224 bp. DNA. linear. PRI 24-FEB-2001
LOCUS            HSJ1059L7
DEFINITION       Human DNA sequence from clone R5-105917 on chromosome
20q13.2-13.33 Contains the 5' end of the TMEMPA1 (PMEPA1) gene
encoding an androgen induced 1b transmembrane protein, ESTs, STSS,
GSSs and two CpG islands, complete sequence.
ACCESSION        AL121913.4 GI:7161781
VERSION          AL121913.4
KEYWORDS         HMG; CpG island; PMEPA1; TMEMPA1; transmembrane protein.
SOURCE           human.
ORGANISM         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE        1 (Bases 1 to 150224)
AUTHORS          Skuue,C.
TITLE            Direct Submission
JOURNAL          Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
COMMENT          On Mar 6, 2000 this sequence version replaced gi:7007305.
During sequence assembly data is compared from overlapping clones.

```

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences without a small overlap as described above.

Only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMEP; Information on the WORMEP database can be found at [http://www.sanger.ac.uk/projects/C\\_elegans/wormep](http://www.sanger.ac.uk/projects/C_elegans/wormep) This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGC/Chr20>

This sequence is the entire insert of clone RP5-1059J7. The true left end of clone RP11-402P1 is at 106677 in this sequence. All regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats. All regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-1059J7 is from the library RCI1-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pcypac2.

```

FEATURES
Source
    Location/Qualifiers
        1..150224
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="q13.2-13.33"
            /clone="RP5-1059L7"
            /clone_1Id="RPCL-5"
            283..355
            /note="MIR repeat: matches 71..145 of consensus"
            1509..1628
            /note="MIR repeat: matches 46..165 of consensus"
            2076..2157
            /note="L2 repeat: matches 2616..2707 of consensus"
            complement(2445..53425)
            /gene="TMEPAI"
            complement(join(<2445..2599,53362..53425))
            /gene="TMEPAI"
            /product="dJ1059L7.1.2 (androgen induced type 1b transmembrane protein (PMEPA1), isoform 2)"
            /note="match: cDNAs: Em:AF242428
            match: ESTs: Em:AA086767"
            /evidence=not_experimental
            complement(join(<2445..2599,53362..53365))
            /gene="TMEPAI"
            /note="continues in dJ718J7 (AL035541)"
            /codon_start=1
            /evidence=not_experimental
            /product="dJ1059L7.1.2 (androgen induced type 1b transmembrane protein (PMEPA1), isoform 2)"
            /protein_id="CAC32857.1"
            /gb_xref="GI:13160408"
            /translation="MALEFVQIIIVVMVVTCLSHYKLSRPTSRHSGC
            RREDLSS"
            complement(join(<2445..2599,52376..52477))
            /gene="TMEPAI"
            /product="dJ1059L7.1.1 (androgen induced type 1b transmembrane protein (PMEPA1), isoform 1)"
            /note="match: cDNAs: Em:AF009426 Em:AF009427 Em:AF009425
            Em:AF009424
            match: ESTs: Em:AA249792 Em:AI594390 Em:AA128075
            Em:AA086767"
            /evidence=not_experimental
            complement(join(<2445..2599,52376..52472))
            /gene="TMEPAI"
            /note="continues in dJ718J7 (AL035541)

```

```
match: proteins: Tr:O15166 Tr:O15168"
/codon_start=1
/evidence=not_experimental
/product="drl059L7.1.1 (androgen induced type 1b transmembrane protein (PMEPA1), isoform 1)"
/protein_id="CA88144.1"
/db_xref="GI:7619746"
/translaton="MGVNSTAAAGQPNVCNCNKRSLFOSMETLEFEVOLIIIVTVMWMMVVITCLLSHYKLSARSLFSRISGQRREDALSS"
repeat_region
  /note="10 copies 3 mer cac 90% conserved"
misc_feature
  /note="match: GSS: Em:A059507"
repeat_region
  /note=".3324 .3324 matches 37 mer 83% conserved"
repeat_region
  /note=".3339 .3413 matches 25 mer 92% conserved"
repeat_region
  /note=".3390 .3921 matches 28 mer 61% conserved"
repeat_region
  /note=".3394 .3893 matches 25 copies 20 mer 59% conserved"
repeat_region
  /note=".3407 .3892 matches 27 copies 18 mer 58% conserved"
repeat_region
  /note=".3421 .3910 matches 7 copies 70 mer 62% conserved"
repeat_region
  /note=".3473 .3904 matches 12 copies 36 mer 58% conserved"
repeat_region
  /note=".3533 .3932 matches 10 copies 38 mer 64% conserved"
repeat_region
  /note=".3573 .3697 matches 5 copies 25 mer 76% conserved"
repeat_region
  /note=".3746 .3921 matches 8 copies 22 mer 64% conserved"
repeat_region
  /note=".4032 .4069 matches 19 copies 2 mer tt 81% conserved"
repeat_region
  /note=".4792 .5001 matches 70 mer 82% conserved"
repeat_region
  /note=".5002 .5190 matches 1.200 of consensus"
repeat_region
  /note=".5224 .5282 matches 94.154 of consensus"
repeat_region
  /note=".8205 .8205 matches 48.161 of consensus"
repeat_region
  /note="MIR repeat: matches 48.161 of consensus"
misc_feature
  /gene="PMEPA1"
  /note="match: GSS: Em:A0881699"
repeat_region
  /note=".9382 .9382 matches 81.239 of consensus"
misc_feature
  /note="MIR repeat: matches 81.239 of consensus"
misc_feature
  /note=".9286 .9706 matches 9295.9861 matches 10035.10140 matches 39.146 of consensus"
repeat_region
  /note="Charlie4 repeat: matches 39.146 of consensus"
repeat_region
  /note=".1115 .1134 matches 11.196 of consensus"
repeat_region
  /note="MIR repeat: matches 11.196 of consensus"
repeat_region
  /note=".11630 .12037 matches 2290.2705 of consensus"
repeat_region
  /note=".12054 .12362 matches 1.311 of consensus"
repeat_region
  /note="AlusX repeat: matches 1.311 of consensus"
repeat_region
  /note=".14448 .14493 matches 73 copies 2 mer at 76% conserved"
repeat_region
  /note=".14634 .14813 matches MER20 repeat: matches 1.187 of consensus"
misc_feature
  /note="MER20 repeat: matches 1.187 of consensus"
repeat_region
  /note="Cpg island"
  /evidence=not_experimental
repeat_region
  /note="MIR repeat: matches 7.254 of consensus"
repeat_region
  /note="MIR repeat: matches 7.254 of consensus"
repeat_region
  /note="MIR repeat: matches 1.374 of consensus"
repeat_region
  /note="MIR repeat: matches 1.374 of consensus"
```

```
misc_feature
/note="L2 repeat: matches 2679. .2736 of consensus"
17699. .18179
/note="match: GSS: Em:HQ03107"
17902. 18039
repeat_region
/note="MIR repeat: matches 107. .243 of consensus"
19070. .19291
repeat_region
/note="MIR repeat: matches 35. .261 of consensus"
19954. .20005
repeat_region
/note="L2 repeat: matches 2358. .2411 of consensus"
20264. .20432
repeat_region
/note="L2 repeat: matches 2581. .2750 of consensus"
20530. .21282
repeat_region
/note="L2 repeat: matches 5595. .6341 of consensus"
21980. .22192
repeat_region
/note="L2 repeat: matches 2322. .2562 of consensus"
23102. .23203
repeat_region
/note="MIR repeat: matches 117. .218 of consensus"
23231. .23438
repeat_region
/note="MER20 repeat: matches 3. .218 of consensus"
23439. .23512
repeat_region
/note="MIR repeat: matches 182. .257 of consensus"
24460. .24580
repeat_region
/note="L2 repeat: matches 6102. .6221 of consensus"
24752. .25042
repeat_region
/note="L2 repeat: matches 2386. .2709 of consensus"
25208. .25295
repeat_region
/note="MIR repeat: matches 61. .155 of consensus"
26450. .26485
repeat_region
/note="18 copies 2 mer by 97% conserved"
29374. .29501
repeat_region
/note="MIR repeat: matches 55. .190 of consensus"
29550. .29728
repeat_region
/note="MIR repeat: matches 69. .248 of consensus"
30974. .31029
```

Query Match 39.6%; Score 522.8; DB 9; Length 150224;  
Best Local Similarity 99.6%; Pred. No. 1.9e-64;  
Matches 524; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 CGACCCGGCTTCGAGAGCAACCGATCTCTTGAGACTTAATGAGAGAGAGAGCGG 60
DB 52896 CGACCCGGCTTCGAGAGCAACCGATCTCTTGAGACTTAATGAGAGAGAGCGG 52837
QY 61 CGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 120
DB 52836 CGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 52777
QY 121 GGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 52776 GGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 52717
QY 181 ACGCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 240
DB 52716 ACGCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 52657
QY 241 TGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 52656 TGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 52597
QY 301 CTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 52596 CTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 52537
QY 361 CCCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 52536 CCCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 52477
QY 421 CTTGATGGGGGTCAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 52476 CTTGATGGGGGTCAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 52417
QY 481 GTGCAACTGCAAAAGCTCTTTGTTCCAGAGCATGAGATCAAGAG 526
|||||
```

```
DB 52416 GTGCAACTGCAAAAGCTCTTTGTTCCAGAGCATGAGATCAAGAG 52371
RESULT 11
AX392430/c 693 bp DNA linear PAT 23-MAR-2002
LOCUS
DEFINITION
Sequence 14 from Patent WO0216416.
ACCESSION
AX392430
VERSION
AX392430.1 GI:19700746
KEYWORDS
human.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and
Turli,T.G.
Diagnosis and treatment of cardiovascular conditions
Patent: WO 0216416-A 14 28-FEB-2002.
JOURNAL
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
FEATURES
source
1. 693
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 101 a 205 c 237 g 149 t 1 others
ORIGIN
Query Match 37.4%; Score 493.6; DB 6; Length 693;
Best Local Similarity 97.4%; Pred. No. 6e-60;
Matches 533; Conservative 0; Mismatches 10; Indels 4; Gaps 3;
QY 776 CCGGCTTCCGGCCAGCGGAGCGCTTCACCGCTTCCAGCCCGCTACCTGACG 835
DB 693 CCGGCTTCCGGCCAGCGGAGCGCTTCACCGCTTCCAGCCCGCTACCTGACG 634
QY 836 CACAGATGACACTGACACACCCACCATCTCGCTGCAGAGGAGAGAGAGCCCGCCACTAC 895
DB 633 CACAGATGCA-CTGCGCCCGCCACCATCTCGCTGCAGAGGAG--AGAGCCCGCCACTAC 577
QY 896 CAGGCGCCCTGACCTTCAGCTTCGGAGCCCGGAGCAGCAGCTGGAACCTGAACCGGGAG 955
DB 576 CAGGCGCCCTGACCTTCAGCTTCGGAGCCCGGAGCAGCAGCTGGAACCTGAACCGGGAG 517
QY 956 TCGGTGGCGCGACCGCCCAAGAAACATCTTTCAGAGTGCAGCTGATGATAGGCCAGG 1015
DB 516 TCGGTGGCGCGACCGCCCAAGAAACATCTTTCAGAGTGCAGCTGATGATAGGCCAGG 457
QY 1016 CTGGGCGCGCCCTGACCGCCCGCCAGCATGTAATCTGGGATCAGCGCCAGCTGACGAGC 1075
DB 456 CTGGGCGCGCCCTGACCGCCCGCCAGCATGTAATCTGGGATCAGCGCCAGCTGACGAGC 397
QY 1076 GGGCGGGCGATGAGAGGGGGCGCGCCCGCCAGCTACAGCGAGGTATCGGCACTACCGGGG 1135
DB 396 GGGCGGGCGATGAGAGGGGGCGCGCCCGCCAGCTACAGCGAGGTATCGGCACTACCGGGG 337
QY 1136 TCGTCTTCCAGACCGCAGCAGAGAGTGGGGCGCGCTCTCTGAGGAG--GGGAGACCGGGCT 1194
DB 336 TCGTCTTCCAGACCGCAGCAGAGAGTGGGGCGCGCTCTCTGAGGAGGGGGAGACCGGGCT 277
QY 1195 CCACACACACATCGCGCCCTAGAGAGCGACCATCTGAGAGCAAAAGAGAGATAA 1254
DB 276 CCACACACACATCGCGCCCTAGAGAGCGACCATCTGAGAGCAAAAGAGAGATAA 217
QY 1255 ACAGAAAGAGACCTCTCTGAGGTCCCGAGAGGGGGCGGGCTGGGGCTGCTGAGGTGAA 1314
DB 216 ACAGAAAGAGACCTCTCTGAGGTCCCGAGAGGGGGCGGGCTGGGGCTGCTGAGGTGAA 157
QY 1315 AAGGAG 1321
DB 156 AAGGAG 150
RESULT 12
```

LOCUS	AF220208	651 bp	mRNA	linear	ROD 05-MAR-2001
DEFINITION	Mus musculus Nedda MW domain-binding protein 4 mRNA, partial cds.				
ACCESSION	AF220208				
VERSION	AF220208.1	GI:12004973			
KEYWORDS					
SOURCE					
ORGANISM	Mus musculus.				
REFERENCE	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 651) Jolliffe,C.N., Harvey,K.F., Haines,B.P., Parasiyram,G. and Kumar,S. Identification of multiple proteins expressed in murine embryos as binding partners for the MW domains of the ubiquitin-protein ligase Nedda Nedda Blochem. J. 351 Pt 3, 557-565 (2000)				
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					

FEATURES	source	1. 812	/organism="Homo sapiens"	/db_xref="taxon:9606"	157 a	272 c	237 g	146 t
BASE COUNT					157 a	272 c	237 g	146 t
ORIGIN								
Query Match		31.6%	Score 417.8;	DB 6;	Length 812;			
Best Local Similarity		77.4%	Pred. No. 2.3e-49;					
Matches 549; Conservative		0;	Mismatches 122;	Indels 38;	Gaps 2			
Y	251	GGCGGCGCGGAACTTGGCGGCGAGCCGAGCCGCGGAGCCGGGCGCGCTCCCGCGCC	310					
Db	130	GTCCTGACAGGCTCTTCATCAGCGGACACAGCAGGAGGAGGAGAGAAAGTGGCTGTC	189					
Y	311	GGCGGCGCTTCGATCGGCGGCGCCAGCTCGGGGCGCGCGAGCCCGCGCGC	370					
Db	190	CTCAGAAAGATGCTGTGGCTCGGAGACAGTGTCAAGGAGAGGATCCACAGAGCGC	249					
Y	371	CCCCGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCATGACCGCTTGATGGG	430					
Db	250	AGCTTACGCGCGCGCTCGGCGACGAGCCGCTGCGGTGGCGCTTCGCGCCAGGAG	309					
Y	431	GTCACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCAATGTCTCTGACGTCAATGC	490					
Db	310	-----CGTTTTCACCGTGTCCAGCCCAATGTCTCTGACCGTCAATGC	355					
Y	491	AAAGCGCTTTGTCACAGAGCATGAGATCAGGAGCTGGAATTTGTTCAATCAATCATC	550					
Db	356	AAAGCGCTTTGTTCCAGAGCATGAGATCAGGAGCTGGAATTTGTTCAATCAATCATC	415					
Y	551	ATCGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	610					
Db	416	ATCGTGTGTG-----TCACGTGCTGCTGAGCCACTACAG	451					
Y	611	CTGCTCAGCGTCTTCATCAGCGCGCACAGCCAGGCGCGGAGAGAGAAATGCTCTG	670					

```

|||||
Db 452 CTGCTCAGGCTCTCTCATACAGCCGACAGCGGCGGAGAGAGAAGATGCCCTG 511
Qy 671 TCTCTAGAAAGATGCTCTGTGGCCCTCGAGAGACATGTTCAGCAACGGAATCCGAG 730
Db 512 TCTCTAGAAAGATGCTCTGTGGCCCTCGAGAGACATGTTCAGCAACGGAATCCGAG 571
Qy 731 CCGAGAGTTCAGGCGCCGCTCGAGAGCGGAGCGGCTGAGCGGCGGCTTGGCCAG 790
Db 572 CCGAGAGTTCAGGCGCCGCTCGAGAGCGGAGCGGCTGAGCGGCGGCTTGGCCAG 631
Qy 791 CCGAGAGCTTCACACCGCTTCACAGCCACCTATCCGTACTGACAGCAGAGATGACTG 850
Db 632 CCGAGAGCTTCACACCGCTTCACAGCCACCTATCCGTACTGACAGCAGAGATGACTG 691
Qy 851 CCACCCACCATCTCTGCTGTAGAGCGGAGAGAGCGCCACCTTACAGAGGCGCTGAC 910
Db 692 CCGCCACCATCTCTGCTGTAGAGCGGAGAGAGCGCCACCTTACAGAGGCGCTGAC 751
Qy 911 CTCACGCTTCGAGCGCCGAGCAGCAGCTGGAACCTGAACCGGAGGCTGG 959
Db 752 TTCAGGCTTCGAGCGCCGAGCAGAGAGTGTGAATTGAACGAGGAGTTGG 800

RESULT 14
AX071267 408 bp DNA linear PAT 25-JAN-2001
LOCUS Sequence 1739 from Patent WO0102568.
DEFINITION AX071267
ACCESSION AX071267
VERSION AX071267.1 GI:12581618
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 408)
WILLIAMS,L.T., ESCOBEDO,J., INIIS,M.A., GARCIA,P.D., KLINGER,J.,
KASSAM,G., REINHARD,C., RANDAZZO,F., KENNEDY,G.C., POT,D.,
LAMSON,G., DRMANAC,R., CRKENTAKOV,R., DRMANAC,S., DICKSON,M.,
LABET,I., LESHKOWITZ,D., KITA,D., GARCIA,V. and STRACHE-CRAIN,B.
Human genes and gene expression products
Patent: WO 0102568-A 1739 11-JAN-2001;
CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
FEATURES
source
1..408
location/Qualifiers
BASE COUNT 81 a 159 c 115 g 53 t
ORIGIN
Query Match 30.4%; Score 401.4; DB 6; Length 408;
Best Local Similarity 99.8%; Pred. No. 5.2e-47;
Matches 402; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 581 GTGATCAGTGCCTGTGTCAGCACTACAGCTGTCTCAGCGTCTTCATCAGCGGCAC 640
Db 6 GAGATCAGTGCCTGTGTCAGCACTACAGCTGTCTCAGCGTCTTCATCAGCGGCAC 65
Qy 641 AGCCAGGGGGGAGAGAGAAGATGCCCTGTCTCAGAGAGATCGTGGCCCTCGAG 700
Db 66 AGCCAGGGGGGAGAGAGAAGATGCCCTGTCTCAGAGAGATCGTGGCCCTCGAG 125
Qy 701 AGCAGACTGTACAGCAAGGAATCCAGAGCGGAGTCTACAGCGCGCGCCGAGCCACC 760
Db 126 AGCAGACTGTACAGCAAGGAATCCAGAGCGGAGTCTACAGCGCGCGCCGAGCCACC 185
Qy 761 GACGCGCTGCGCGCTGCGCGCTTCGCGAGGAGCGCTTCACACCGCTTCACGCCACC 820
Db 186 GACGCGCTGCGCGCTGCGCGCTTCGCGAGGAGCGCTTCACACCGCTTCACGCCACC 245
Qy 821 TATCCGTAAGTTCAGACAGAGATGACTGACACCCACCATCTGCTGTACAGCGGAG 880
Db 246 TATCCGTAAGTTCAGACAGAGATGACTGACACCCACCATCTGCTGTACAGCGGAG 305

```

```

Qy 881 GAGCCCCCACCCTACACAGGCGCCCTGCACCTCCAGCTTGGGAGCCCGAGCAGACTG 940
Db 306 GAGCCCCCACCCTACACAGGCGCCCTGCACCTCCAGCTTGGGAGCCCGAGCAGACTG 365
Qy 941 GAACTGAACCGGAGTGTGTCGCGCACCCCAACCAAGAACCA 983
Db 366 GAACTGAACCGGAGTGTGTCGCGCACCCCAACCAAGAACCA 408

RESULT 15
AL837509 156075 bp DNA linear HTG 17-AUG-2002
LOCUS Mus musculus chromosome 2 clone RP23-44L6, *** SEQUENCING IN
DEFINITION PROGRESS ***, 4 unordered pieces.
ACCESSION AL837509
VERSION AL837509.4 GI:22416197
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 156075)
Almeida,J.
Direct Submission
Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Aug 21, 2002 this sequence version replaced gi:22204793.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: bm4416
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Consensus quality: 155649 bases at least Q40
Consensus quality: 155716 bases at least Q30
Consensus quality: 155743 bases at least Q20
Insert size: 155775; sum-of-contigs
Insert size: 162242; 2.6% error; agarose-fp
Quality coverage: 7.64x in Q20 bases; sum-of-contigs quality
coverage: 7.40x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 9981: contig of 9981 bp in length
* 9982 10081: gap of 100 bp
* 10082 13090: contig of 3009 bp in length
* 13091 13190: gap of 100 bp
* 13191 95958: contig of 82768 bp in length
* 95959 96058: gap of 100 bp
* 96059 156075: contig of 60017 bp in length.
location/Qualifiers
1..156075
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-44L6"
/clone_1lb="RPCT-23"
1..9981
/note="assembly_fragment:00303
fragment_chain:1"
10082..13090
misc_feature
misc_feature
misc_feature

```





PA (BGHM) BRIGHAM & WOMEN'S HOSPITAL INC.  
 PA (PFIZ) PFIZER INC.  
 XX  
 PI Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;  
 XX WPI, 2002-280912/32.  
 DR P-PSDB; AAU78231.  
 XX

PT Novel nucleic acid molecule encoding Mechanically Induced Vascular  
 Receptor-1 polypeptide, useful for treating cardiovascular diseases  
 XX  
 PS Claim 2; Page 87-88; 105pp; English.

XX  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having a  
 CC cardiac cell anti-apoptotic activity and fragments of it provided  
 CC they are not identical to Genbank sequences A161441.1, A1594390,  
 CC NM\_004338 and A0177461. Also included are expression vectors, host  
 CC cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of  
 CC MIVR-1, contacting a molecule having cardiac cell anti-apoptotic activity  
 CC with a candidate agent, where the molecule is a nucleic acid molecule  
 CC comprising MIVR-1, IEX-1, VDUP-1, BFG-2 and FIS-110 or its  
 CC expression product, determining if the anti-apoptotic activity is  
 CC modulated and thereby identifying a modulator. The cardiac cell anti-  
 CC apoptotic molecules and nucleic acids of the invention are useful for  
 CC treating, diagnosing and monitoring progression of such diseases and  
 CC disorders as characterised by increased apoptotic cell-death of vascular  
 CC endothelial cells e.g. cardiac hypertrophy, myocardial infarction,  
 CC stroke, arteriosclerosis and heart failure. The present sequence  
 CC encodes human MIVR-1.  
 XX

SQ Sequence 1321 BP; 223 A; 493 C; 440 G; 165 T; 0 other;

Query Match 100.0%; Score 1321; DB 24; Length 1321;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-203;  
 Matches 1321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACCGCGGTCTCGAGCAAAACCGATCTCTTGACTTGAATGAGAGAGAGCGG 60  
 Db 1 CGACCGCGGTCTCGAGCAAAACCGATCTCTTGACTTGAATGAGAGAGAGCGG 60  
 QY 61 CGCGCGCGCGCGCGCGCGCGCTGCGTGGGAAAGCTAGCGGAGAGCTACGCCG 120  
 Db 61 CGCGCGCGCGCGCGCGCGCGCTGCGTGGGAAAGCTAGCGGAGAGCTACGCCG 120  
 QY 121 GCGCGAGCGCGCGCGCGCGCTGCGAAGCCATTTCGAGCGGACCGCGGCGACAC 180  
 Db 121 GCGCGAGCGCGCGCGCGCGCTGCGAAGCCATTTCGAGCGGACCGCGGCGACAC 180  
 QY 181 AGCGCGCGCGCGCGCGCGCGCGCGAGCGCGGAGCGGAGCGGCGGCGCGAC 240  
 Db 181 AGCGCGCGCGCGCGCGCGCGCGCGAGCGCGGAGCGGAGCGGCGGCGCGAC 240  
 QY 241 TGAGCGCGCGCGCGCGCGCGCGCGGAACTTGCGGAGCGGCGGCGGCGCGC 300  
 Db 241 TGAGCGCGCGCGCGCGCGCGCGCGGAACTTGCGGAGCGGCGGCGGCGCGC 300  
 QY 301 CTCCCCCGCGCGCGCGCTGCTGATGCGGCGCGCGAGCTCGGCGGCGCGGAGCC 360  
 Db 301 CTCCCCCGCGCGCGCGCTGCTGATGCGGCGCGCGAGCTCGGCGGCGCGGAGCC 360  
 QY 361 CCCGC 420  
 Db 361 CCCGC 420  
 QY 421 CTTGATGGGGGTCAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 480  
 Db 421 CTTGATGGGGGTCAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 480  
 QY 481 GTGCAACTGCAAAAGCTCTTTGTTCAGAGCATGAGATCAGGAGCTGAGATTGTTC 540  
 Db 481 GTGCAACTGCAAAAGCTCTTTGTTCAGAGCATGAGATCAGGAGCTGAGATTGTTC 540

QY 541 GATCATCATCATGTCGT 600  
 Db 541 GATCATCATCATGTCGT 600  
 QY 601 CCAGTACAAGCTGTGTGACAGGCTCTTCATCAGCGCGCACAGCCAGCGCGGAGAGAGA 660  
 Db 601 CCAGTACAAGCTGTGTGACAGGCTCTTCATCAGCGCGCACAGCCAGCGCGGAGAGAGA 660  
 QY 661 AGATGCGCTGTCTCTAGAAAGATGCTGTGCGCTTCGAGACACAGTTCAGGACCGG 720  
 Db 661 AGATGCGCTGTCTCTAGAAAGATGCTGTGCGCTTCGAGACACAGTTCAGGACCGG 720  
 QY 721 AATCCAGAGCGCGAGGTATAGCCCGCGCTGCGCGCACCGCGCGGCGGTGCGCGCC 780  
 Db 721 AATCCAGAGCGCGAGGTATAGCCCGCGCTGCGCGCACCGCGCGGCGGTGCGCGCC 780  
 QY 781 CTTGCGCGCGCGGAGCGGCTTCACCGCTTCAGCGCCACACTTCGCTACCTGACAGAGA 840  
 Db 781 CTTGCGCGCGGAGCGGCTTCACCGCTTCAGCGCCACACTTCGCTACCTGACAGAGA 840  
 QY 841 GATCGACTGCGCACCCACCATCTCGCTGTCAAGCGGAGAGAGCCCGACCTACCAAGG 900  
 Db 841 GATCGACTGCGCACCCACCATCTCGCTGTCAAGCGGAGAGAGCCCGACCTACCAAGG 900  
 QY 901 CCCCTGCAACCTTCAGCTTCGCGGACCGCGAGCAGCTGGAATGAACCGGAGTGGT 960  
 Db 901 CCCCTGCAACCTTCAGCTTCGCGGACCGCGAGCAGCTGGAATGAACCGGAGTGGT 960  
 QY 961 GCGCGACCCCGCAACAGAACCATCTTGACAGTGAATGATGATGATGATGATGATGAT 1020  
 Db 961 GCGCGACCCCGCAACAGAACCATCTTGACAGTGAATGATGATGATGATGATGATGAT 1020  
 QY 1021 GCGCGCGCGCGCGCGCGCGAGTACGCGGCGATCGGCGCGCGCGCGCGCGCGCG 1080  
 Db 1021 GCGCGCGCGCGCGCGCGCGAGTACGCGGCGATCGGCGCGCGCGCGCGCGCGCG 1080  
 QY 1081 GCGCATGAGAGGCG 1140  
 Db 1081 GCGCATGAGAGGCG 1140  
 QY 1141 CTTCCAGCACAGCAGAGAGAGTGGGCGCGCTCTTGCTGTGAGGAGACCGGCTCCACCA 1200  
 Db 1141 CTTCCAGCACAGCAGAGAGAGTGGGCGCGCTCTTGCTGTGAGGAGACCGGCTCCACCA 1200  
 QY 1201 CACACACATCG 1260  
 Db 1201 CACACACATCG 1260  
 QY 1261 AGGACACCTCTCTAGGGTCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320  
 Db 1261 AGGACACCTCTCTAGGGTCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320  
 QY 1321 G 1321  
 Db 1321 G 1321

RESULT 2  
 AAA75151  
 ID AAA75151 standard; cDNA: 969 BP.  
 XX  
 AC AAA75151;  
 XX  
 DT 15-JAN-2001 (first entry)  
 XX  
 DE cDNA encoding a human TANGO 261 polypeptide.  
 XX  
 XX TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;



KW	cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW	brain herniation; iatrogenic disease; inflammation; meningitis;
KW	Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;
KW	multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
OS	Homo sapiens.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	6..764
FT	/tag= a
FT	/product= "TANGO 261"
FT	sig_peptide
FT	6..89
FT	/tag= b
FT	mat_peptide
FT	90..764
FT	/tag= c
XX	
FN	WO200052022-A1.
PD	
XX	
XX	08-SEP-2000.
XX	
XX	01-MAR-2000; 2000WO-US05226.
XX	
XX	01-MAR-1999; 99US-0122458.
XX	
PA	(MILL-) MILLERINUM PHARM INC.
XX	
PI	Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
DR	WPI: 2000-579269/54.
DR	P-PSDB: AAB18449.
XX	
PT	Novel human and murine secreted proteins designated TANGO 216, 261,
PT	262, 266 and 267 useful as modulating agents of cellular processes,
PT	e.g. for treating cancer -
PS	Claim 2; Fig 5; 175pp; English.
XX	
XX	The present sequence encodes a human TANGO 261 polypeptide. The
CC	specification also describes TANGO 266, TANGO 216, TANGO 262, and
CC	TANGO 267. The TANGO polypeptides can be used to modulate cellular
CC	proliferation, modulate cellular differentiation and/or modulate
CC	cellular adhesion. The proteins can be used to treat any von Willebrand
CC	factor-associated disorder, regulate extracellular matrix structuring,
CC	cellular adhesion, and cell trafficking and/or migration, modulate
CC	cellular interactions, modulate cell adhesion in proliferative
CC	disorders, such as cancer, modulate the proliferation, differentiation,
CC	and/or function of cells that appear in the bone marrow, and leukocytes,
CC	treat bone marrow, blood and hematopoietic associated diseases and
CC	disorders, atelectasis, pulmonary congestion or oedema, emphysema,
CC	chronic bronchitis, bronchial asthma and bronchiectasis, intestinal
CC	disorders, spleen associated diseases, modulate renal disorders, treat
CC	cardiovascular disorders such as ischemic heart disease, modulate the
CC	proliferation, differentiation, and/or function of bone and cartilage
CC	cells and to treat bone and/or cartilage associated diseases or
CC	disorder. They may also be used to treat disorders associated with the
CC	ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic
CC	disease, inflammatory, bacterial and viral meningitis, Alzheimer's
CC	disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,
CC	brain cancers, hydrocephalus and encephalitis, and treat hepatic
CC	disorders.
XX	
XX	Sequence 969 BP; 211 A; 316 C; 294 G; 148 T; 0 other;
SO	
XX	
XX	Query Match 60.8%; Score 803.2; DB 21; Length 969;
XX	Best Local Similarity 99.6%; Pred. No. 1.5e-120;
XX	Matches 805; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX	
OY	514 GGAGATCAGCAGAGCTGGAGTTGTCAGATCATCATCATGTCGTCGATGATGAT 573
DB	
DB	2 GGAGATGGGGGAGCTGGAGTTGTCAGATCATCATCATCGTGGTGTATGATGTAT 61
OY	574 GGTGTGTGATACGCTGCTCTGACGCATACAGCTGTGTGACGCTCTTATCATCG 633

Db	62	GGTGTGGTATATCACTGTCGCTGCTGAGCCACTACAAAGCTGTCTGCACGSGTCTTCATACG	121
QY	634	CCGGACACGCCCCGGGGGGAGAGAGAATGCCCCCTGTCTCTCAGAAAGATGCTTGTGCC	693
Db	122	CCGGACACGCCCCGGGGGGAGAGAGAATGCCCCCTGTCTCTCAGAAAGATGCTTGTGCC	181
QY	694	CTCGGAGAGCACATGTTCAGGCAACGGAATCCAGAGCCCGCAGGTCTACGCCCGCTTCG	753
Db	182	CTCGGAGAGCACATGTTCAGGCAACGGAATCCAGAGCCCGCAGGTCTACGCCCGCTTCG	241
QY	754	GCCCCACGACCCGCTTGGCCCGGGCCGCTTCCGCCAGCGGAGGAGCGCTTCACCGCTTCA	813
Db	242	GCCCCACGACCCGCTTGGCCCGGGCCGCTTCCGCCAGCGGAGGAGCGCTTCCACCGCTTCA	301
QY	814	GCCCCACCTATCCCTGACCTGCGAGCAGCAGATGCACCTGCGACCCACCACCATCTCGCTGTAGA	873
Db	302	GCCCCACCTATCCCTGACCTGCGAGCAGCAGATGCACCTGCGCGCCACACCATCTCGCTGTAGA	361
QY	874	CGGGGAGGAGCCCCCACCCTTACACAGGGGCCCTTCACACCTTCACGCTTCGGGAGCCCGACGA	933
Db	362	CGGGGAGGAGCCCCCACCCTTACACAGGGGCCCTTCACACCTTCACGCTTCGGGAGCCCGACGA	421
QY	934	GCAGCTGGAATGSAACCGGGGGGTGCGGGCGCGACCCGCAACCAAGAACCATCTTGACAG	993
Db	422	GCAGCTGGAATGSAACCGGGGGGTGCGGGCGCGACCCGCAACCAAGAACCATCTTGACAG	481
QY	994	TGACCTGATGATGTAGTCCACAGGCTGCGGGCGGCCCTTCGCCCCCGACAGATTAACCTCGGCAT	1053
Db	482	TGACCTGATGATGTAGTCCACAGGCTGCGGGCGGCCCTTCGCCCCCGACAGATTAACCTCGGCAT	541
QY	1054	CAGCGCCACGCTGTACGCGCAGCGCGGGGGCGCATGAGAGGGGCGCGGCCACCTTACAGCGA	1113
Db	542	CAGCGCCACGCTGTACGCGCAGCGCGGGGGCGCATGAGAGGGGCGCGGCCACCTTACAGCGA	601
QY	1114	GGCTATGGGGCCACTACCCGGGGGTCTCTCTTCAGACACACAGAGAGAGTGGGGCGGCCCTC	1173
Db	602	GGCTATGGGGCCACTACCCGGGGGTCTCTCTTCAGACACACAGAGAGAGTGGGGCGGCCCTC	661
QY	1174	CTTGCTGGAGGGGACCCGGGCTCCACACACACACATCGCGCCCTTAGAGAGGCGCAGCCAT	1233
Db	662	CTTGCTGGAGGGGACCCGGGCTCCACACACACACATCGCGCCCTTAGAGAGGCGCAGCCAT	721
QY	1234	CTGAGGACAAAGAGAGATTAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGCGG	1293
Db	722	CTGAGGACAAAGAGAGATTAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGCGG	781
QY	1294	GGCTGGGGCTGCCCTTAGAGTGAAAAAGCGAG	1321
Db	782	GGCTGGGGCTGCCCTTAGAGTGAAAAAGCGAG	809
RESULT 3			
ABK92120			
ID	ABK92120 standard; DNA: 1140 BP.		
XX	ABK92120;		
AC	15-AUG-2002 (first entry)		
XX	Prostate cancer-associated DNA sequence #6.		
DE	Prostate cancer-associated DNA sequence #6.		
XX	Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;		
KW	gene therapy; gene; ds.		
XX	Mammalia.		
OS	Mammalia.		
XX	Mammalia.		
PN	M0200230268-A2.		
XX	18-APR-2002.		
PD	12-OCT-2001; 2001MO-US32045.		
XX	13-OCT-2000; 2000US-0687576.		
PR			



DR P-PSDB; AAB18461.  
 XX Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 XX  
 PS Disclosure; Page -: 175pp; English.  
 XX  
 CC AAA75163-65 encode human TANGO 261 proteins. The specification also  
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat any von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
 CC and cell trafficking and/or migration, modulate cellular interactions,  
 CC modulate cell adhesion in proliferative disorders, such as cancer,  
 CC modulate the proliferation, differentiation, and/or function of cells  
 CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 CC and hematopoietic associated diseases and disorders, atelectasis,  
 CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, spleen associated  
 CC diseases, modulate renal disorders, treat cardiovascular disorders such  
 CC as ischemic heart disease, modulate the proliferation, differentiation,  
 CC and/or function of bone and cartilage cells and to treat bone and/or  
 CC cartilage associated diseases or disorder. They may also be used to  
 CC treat disorders associated with the ovaries, and cerebral oedema,  
 CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
 CC bacterial and viral meningitis, Alzheimer's disease, cerebral  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC hydrocephalus and encephalitis, and treat hepatic disorders.  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided.  
 CC  
 XX  
 SO Sequence 969 BP: 211 A; 317 C; 293 G; 148 T; 0 other;

Query Match 60.7%; Score 801.6; DB 21; Length 969;  
 Best Local Similarity 99.5%; Pred. No. 2.8e-120;  
 Matches 804; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 514 GGAATATCAGGAGCTGAGATTGTTTCAGATCATCATCTGTTGATGATGATGAT 573  
 DB 2 GGAATATCAGGAGCTGAGATTGTTTCAGATCATCATCTGTTGATGATGATGAT 61  
 QY 574 GGTGGTGTGATCAGTGCCTGCTGAGCACTACAGTGTCTCAGAGGTCCTCATCAG 633  
 DB 62 GGTGGTGTGATCAGTGCCTGCTGAGCACTACAGTGTCTCAGAGGTCCTCATCAG 121  
 QY 634 CCGGACACAGCCAGGGGGGAGAGAGAGATGCTGTCTCAGAGAGATGCTGTGGCC 693  
 DB 122 CCGGACACAGCCAGGGGGGAGAGAGAGATGCTGTCTCAGAGAGATGCTGTGGCC 181  
 QY 694 CTGGAGAGCACTGTTCAGCAAGCAATCCAGAGCCGAGGTCTACGCCGCCCTCG 753  
 DB 182 CTGGAGAGCACTGTTCAGCAAGCAATCCAGAGCCGAGGTCTACGCCGCCCTCG 241  
 QY 754 GCCACAGGACCGCTGGCGGCTGCGCCCTTGCCAGGGAGGGAGGCTTCCAGCCCTTCCA 813  
 DB 242 GCCACAGGACCGCTGGCGGCTGCGCCCTTGCCAGGGAGGGAGGCTTCCAGCCCTTCCA 301  
 QY 814 GCCACCTATCCGTAAGTGCAGCAGAGATGACCTGACCCACCACTCTGCTGTGAGA 873  
 DB 302 GCCACCTATCCGTAAGTGCAGCAGAGATGACCTGACCCACCACTCTGCTGTGAGA 361  
 QY 874 CGGGAGAGAGCCCGACCTTACAGGGCCCTGCACTCTTGAGAGCCCGGAGCA 933  
 DB 362 CGGGAGAGAGCCCGACCTTACAGGGCCCTGCACTCTTGAGAGCCCGGAGCA 421  
 QY 934 GCACCTGGAATGAACCGGAGTGGTGGCGGACCCCAACAAAGAACCACTTTCAGAG 993  
 DB 422 GCACCTGGAATGAACCGGAGTGGTGGCGGACCCCAACAAAGAACCACTTTCAGAG 481  
 QY 994 TGACCTGATGATAGTGCAGAGCTGGGGGGCCCTGCCCCCAGAGTAATCTCGGGCAT 1053  
 DB 1053 TGACCTGATGATAGTGCAGAGCTGGGGGGCCCTGCCCCCAGAGTAATCTCGGGCAT 1053

DB 482 TGACCTGATGATAGTGCAGAGCTGGGGGGCCCTGCCCCCAGAGTAATCTCGGGCAT 541  
 QY 1054 CAGCGCCACGCGTACAGGAGCGGCGGCGCATGAGAGGGCGCGCCCACTACAGCA 1113  
 DB 542 CAGCGCCACGCGTACAGGAGCGGCGGCGCATGAGAGGGCGCGCCCACTACAGCA 601  
 QY 1114 GGTATGCGGCGTACAGGAGCGGCGGCGCATGAGAGGGCGCGCCCACTACAGCA 1173  
 DB 602 GGTATGCGGCGTACAGGAGCGGCGGCGCATGAGAGGGCGCGCCCACTACAGCA 661  
 QY 1174 CTGCTGAGAGGGAGCCCGGCTCCACACACACATGCGCCCTAGAGAGCGAGCAT 1233  
 DB 662 CTGCTGAGAGGGAGCCCGGCTCCACACACACATGCGCCCTAGAGAGCGAGCAT 721  
 QY 1234 CTGAGCAAGAGAGATTAACAGAAAGACACCTCTAGGCTCCAGAGGGGGCGG 1293  
 DB 722 CTGAGCAAGAGAGATTAACAGAAAGACACCTCTAGGCTCCAGAGGGGGCGG 781  
 QY 1294 GCGTGGGCTGCTAGGTGAAAAGCAG 1321  
 DB 782 GCGTGGGCTGCTAGGTGAAAAGCAG 809

RESULT 5  
 ID AAA75164 standard; cDNA; 969 BP.  
 XX  
 AC AAA75164;  
 XX  
 DT 15-JAN-2001 (first entry)  
 XX  
 DE cDNA clone encoding a human TANGO 261 polypeptide.  
 KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; se.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 6..764  
 FT /tag= a  
 FT /product= "TANGO 261"  
 XX  
 PN WO200052022-A1.  
 XX  
 PD 08-SEP-2000.  
 XX  
 PF 01-MAR-2000; 2000WO-US05226.  
 XX  
 PR 01-MAR-1999; 99US-0122458.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
 DR WPI; 2000-579269/54.  
 DR P-PSDB; AAB18462.  
 XX  
 PT Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 XX  
 PS Disclosure; Page -: 175pp; English.  
 XX  
 CC AAA75163-65 encode human TANGO 261 proteins. The specification also

describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO polypeptides can be used to modulate cellular proliferation, modulate cellular differentiation and/or modulate cellular adhesion. The proteins can be used to treat any von Willebrand factor-associated disorder, regulate extracellular matrix structuring, cellular adhesion, and cell trafficking and/or migration, modulate cellular interactions, modulate cell adhesion in proliferative disorders, such as cancer, modulate the proliferation, differentiation, and/or function of cells that appear in the bone marrow, and leukocytes, treat bone marrow, blood and hematopoietic associated diseases and disorders, atelectasis, bronchial asthma and bronchiectasis, intestinal disorders, spleen associated diseases, modulate renal disorders, treat cardiovascular disorders such as ischemic heart disease, modulate the proliferation, differentiation, and/or function of bone and cartilage cells and to treat bone and/or cartilage associated diseases or disorder. They may also be used to treat disorders associated with the ovaries, and cerebral oedema, hydrocephalus, brain herniations, iatrogenic disease, inflammations, bacterial and viral meningitis, Alzheimer's disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers, hydrocephalus and encephalitis, and treat hepatic disorders. note: the present sequence does not appear in the specification: it was created using information provided.

Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other;

Query Match 60.7%; Score 801.6; DB 21; Length 969;  
Best Local Similarity 99.5%; Pred. No. 2.8e-120;  
Matches 804; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

0Y 514 GGAGATCAGCAGAGCTGAGTTGTTTCAGATCATCATCTGCTGATGATGATGAT 573  
DB 2 GGAGATGCGGAGCTGAGTTGTTTCAGATCATCATCTGCTGATGATGATGATGAT 61  
0Y 574 GGTGTGTGTATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633  
DB 62 GGTGTGTGTATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121  
0Y 634 CCGGACAGCAGGAGGCGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 693  
DB 122 CCGGACAGCAGGAGGCGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181  
0Y 694 CTGGAGAGCAGAGTGTAGGCAACGGAATCCAGAGCCGAGGTCTAGCCGCTCTCG 753  
DB 182 CTGGAGAGCAGAGTGTAGGCAACGGAATCCAGAGCCGAGGTCTAGCCGCTCTCG 241  
0Y 754 GCCCACCAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 813  
DB 242 GCCCACCAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 301  
0Y 814 GCCCACCAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 873  
DB 302 GCCCACCAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 361  
0Y 874 CGGGAGAGGAGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 933  
DB 362 CGGGAGAGGAGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 421  
0Y 934 GCAGCTGGAAGTGAACCGGAGTGGTGGCGGACCCCAACAGACATCTTGGACAG 993  
DB 422 GCAGCTGGAAGTGAACCGGAGTGGTGGCGGACCCCAACAGACATCTTGGACAG 481  
0Y 994 TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053  
DB 482 TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541  
0Y 1054 CAGGCGCAGCTGCTAGCGCAGCGGCGGCGATGAGAGGCGGCGGCGGCGGCGGCGG 1113  
DB 542 CAGGCGCAGCTGCTAGCGCAGCGGCGGCGATGAGAGGCGGCGGCGGCGGCGGCGG 601  
0Y 1114 GGTATGCGGCACTACCGGGGGTCTCTTCACAGCAGCAGAGCAGTGGGCGGCGGCTC 1173  
DB 602 GGTATGCGGCACTACCGGGGGTCTCTTCACAGCAGCAGAGCAGTGGGCGGCGGCTC 661

0Y 1174 CTGTGTGAGGGGAGCCGGCTGCACACACATATGCGGCGGCTAGAGAGCGAGCCAT 1233  
DB 662 CTGTGTGAGGGGAGCCGGCTGCACACACATATGCGGCGGCTAGAGAGCGAGCCAT 721  
0Y 1234 CTGGAGCAAGAAGAGTAAACAGAAAGAGACACCTCTTAGGGTCCCGAGGGGGCGG 1293  
DB 722 CTGGAGCAAGAAGAGTAAACAGAAAGAGACACCTCTTAGGGTCCCGAGGGGGCGG 781  
0Y 1294 GCGTGGGCGCTGCTAGCTGAGTGAAGAGCAG 1321  
DB 782 GCGTGGGCGCTGCTAGCTGAGTGAAGAGCAG 809  
RESULT 6  
AAAT75165  
ID AAAT75165 standard; cDNA: 969 BP.  
AC AAAT75165;  
XX  
DT 15-JAN-2001 (first entry)  
DE cDNA clone encoding a human TANGO 261 polypeptide.  
XX  
KW TANGO 266; TANGO 261; TANGO 262; TANGO 267;  
KW cellular proliferation; cellular differentiation; cell trafficking; cancer;  
KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
KW intestinal disorder; spleen associated disease; renal disorder;  
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
KW brain herniation; iatrogenic disease; inflammation; meningitis;  
KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;  
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 6..764  
FT /tag= a  
FT /product= "TANGO 261"  
XX  
XX W0200052022-A1.  
XX  
XX 08-SEP-2000.  
XX  
XX 01-MAR-2000; 2000WO-US05226.  
XX  
XX 01-MAR-1999; 99US-0122458.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
XX  
XX WPI: 2000-579269/54.  
XX P-PSDB: AAB18463.  
XX  
XX Novel human and murine secreted proteins designated TANGO 216, 261,  
XX 262, 266 and 267 useful as modulating agents of cellular processes,  
XX e.g. for treating cancer -  
XX  
XX Disclosure; Page -: 175pp; English.  
XX  
XX AAAT75163-65 encode human TANGO 261 proteins. The specification also  
XX describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
XX polypeptides can be used to modulate cellular proliferation, modulate  
XX cellular differentiation and/or modulate cellular adhesion. The  
XX proteins can be used to treat any von Willebrand factor-associated  
XX disorder, regulate extracellular matrix structuring, cellular adhesion,  
XX and cell trafficking and/or migration, modulate cellular interactions,  
XX modulate cell adhesion in proliferative disorders, such as cancer,  
XX modulate the proliferation, differentiation, and/or function of cells  
XX that appear in the bone marrow, and leukocytes, treat bone marrow, blood

CC and hematopoietic associated diseases and disorders, atelectasis,  
 CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, spleen associated  
 CC diseases, modulate renal disorders, treat cardiovascular disorders such  
 CC as ischemic heart disease, modulate the proliferation, differentiation,  
 CC and/or function of bone and cartilage cells and to treat bone and/or  
 CC cartilage associated diseases or disorder. They may also be used to  
 CC treat disorders associated with the ovaries, and cerebral oedema,  
 CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
 CC bacterial and viral meningitis, Alzheimer's disease, cerebral  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC hydrocephalus and encephalitis, and treat hepatic disorders.  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided.

XX Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other;

Query Match 60.7%; Score 801.6; DB 21; Length 969;

Best Local Similarity 99.5%; Pred. No. 2.8e-120;

Matches 804; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

DB 514 GGAGATACGAGCTGGATTGTTTCAGATCATCATCGTGGTGTGATGATGAT 573  
 2. GGAGATGCGGAGCTGAGATTGTTTCAGATCATCATCGTGGTGTGATGATGAT 61  
 QY 574 GGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 633  
 62 GGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121  
 DB 634 CCGGACAGCCAGGCGGAGAGAGAGAGATGCTCTCTCAGAGAGATGCTGTG 693  
 122 CCGGACAGCCAGGCGGAGAGAGAGAGATGCTCTCTCAGAGAGATGCTGTG 181  
 DB 694 CTGCGAAGACATGTCAGGCAACGGAATCCAGAGCCGAGTCTTACGCCCT 753  
 182 CTGCGAAGACATGTCAGGCAACGGAATCCAGAGCCGAGTCTTACGCCCT 241  
 QY 754 GCCCAGCAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCT 813  
 242 GCCCAGCAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCT 301  
 DB 814 GCCCAGCAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCT 873  
 302 GCCCAGCAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCT 361  
 QY 874 CCGGAGAGAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCT 933  
 362 CCGGAGAGAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCT 421  
 DB 934 GCAGCTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 993  
 422 GCAGCTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481  
 QY 994 TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053  
 482 TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541  
 DB 1054 CAGGCGCAGCTGTCAGGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1113  
 542 CAGGCGCAGCTGTCAGGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601  
 QY 1114 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173  
 602 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661  
 DB 1174 CTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1233  
 662 CTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 721  
 QY 1234 CTGAGAGCAAG 1293  
 722 CTGAGAGCAAG 781

QY 1294 GGCTGGGCTGCTAGGTGAGAAAGCAG 1321  
 DB 782 GGCTGGGCTGCTAGGTGAGAAAGCAG 809

RESULT 7  
 ID AAAA7429  
 AC AAAA7429 standard; DNA: 1061 BP.  
 XX  
 AC AAAA7429;  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX

DE Sequence encoding human neuron-associated protein.

KW Neuron associated protein; NEUP; neurological disorder; epilepsy;  
 KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
 KW Alzheimer's disease; Pick's disease; Huntington's disease;  
 KW dementia; Parkinson's disease; demyelinating disease; meningitis;  
 KW prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;  
 KW cerebral palsy; muscular dystrophy; central nervous system; CNS;  
 KW peripheral nervous system; PNS; myopathy; schizophrenia;  
 KW actinic keratosis; arteriosclerosis; atherosclerosis; bursitis;  
 KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;  
 KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
 KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
 KW AIDS; Addison's disease; adult respiratory distress syndrome;  
 KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
 KW Werner syndrome; trauma; human; ds.  
 XX  
 OS Homo sapiens.  
 XX

FH Key location/Qualifiers  
 FT 101..859  
 FT CDS /\*tag= a  
 FT /product= Neuron associated protein

WO200034477-A2.

15-JUN-2000.

10-DEC-1999; 99MO-US30408.

11-DEC-1998; 98US-0210083.

11-DEC-1998; 98US-9123456.

09-FEB-1999; 99US-0119365.

16-MAR-1999; 99US-0124687.

(INCY-) INCYTE PHARM INC.

PI Tang YT, Yue H, Baughn MR, Hillman JL, Lai P, Au-young J, Yang J;  
 PI Lu DM, Azimzal Y;

DR WPI: 2000-423423/36.  
 DR P-PSDB: AAB01388.

PT New human neuron-associated proteins and polynucleotides encoding them,  
 PT useful for diagnosis, treatment and prevention of cell proliferative  
 PT disorders including cancer, neuronal and neurological disorders  
 PS Claim 9; Page 136; 145pp; English.

CC Human neuron-associated proteins (NEUP) can be used for for  
 CC treating or preventing a disorder associated with decreased  
 CC expression or activity of NEUP. Antagonists of NEUP are useful for  
 CC treating or preventing disorder associated with increased expression  
 CC or activity of NEUP. NEUP or their fragments or derivatives are  
 CC useful for treating neurological disorder such as epilepsy. Ischemic  
 CC cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's  
 CC disease, Pick's disease, Huntington's disease, dementia and  
 CC Parkinson's disease. NEUPs are also useful for treating other  
 CC demyelinating diseases, bacterial and viral meningitis, prion  
 CC diseases including kuru, Creutzfeldt-Jakob disease, nutritional and

0

Qy	1219	AGAGACCGGAGCCATCTGTGAGCAGAGAGATTAACAGAAAGACACCCCTGCTAGGG	1278
Db	802	AGAGAGCGCAGCCATCTGTGAGCAGAGAGATTAACAGAAAGACACCCCTGCTAGGG	861
Qy	1279	TCGCCAGGGGGCGCGGCTGGGGGCTGCCTAGGTGAAAAAGGCAG	1321
Db	862	TCGCCAGGGGGCGCGGCTGGGGGCTGCCTAGGTGAAAAAGGCAG	904
RESULT 8			
ID	AA157868	standard; cDNA; 1066 BP.	
XX			
AC	AA157868;		
XX			
DT	22-OCT-2001	(first entry)	
XX			
DE	Human polynucleotide SEQ ID NO 71.		
XX			
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukaemia; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200153312-A1.		
XX			
PD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000WO-US34263.		
XX			
PR	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou P, Goodrich R, Dirmann RT;		
XX			
XX	WPI: 2001-442253/47.		
DR	P-PSDB; AAM38712.		
XX			
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
XX	such as central nervous system injuries -		
PS	Claim 1; SEQ ID NO 71; 10078bp; English.		
CC			
CC	The invention relates to human nucleic acids (AA157798-AA161369) and		
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities, chemotactic/chemokinetic activity, haemostatic		
CC	activity/inhibin activity, chemotactic/chemokinetic activity, haemostatic		
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S disorders.		
CC	Note: The sequence data for this patent did not form part of the printed		













PD		08-SEP-2000.
XX	PF	01-MAR-2000; 200OWO-US05226.
XX	PR	01-MAR-1999; 99US-0122458.
XX	PA	(MILL-) MILLENNIUM PHARM INC.
XX	P1	Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
XX	DR	WPI; 2000-579269/54.
XX	P-PSDB:	AAB18464.
PT		Novel human and murine secreted proteins designated TANGO 216, 261,
XX		262, 266 and 267 useful as modulating agents of cellular processes,
PS		e.g. for treating cancer -
XX		Disclosure: Page -: 175pp; English.
CC		AAA/5166-68 encode murine TANGO 261 proteins. The specification also
CC		describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC		polypeptides can be used to modulate cellular proliferation, modulate
CC		cellular differentiation and/or modulate cellular adhesion. The
CC		proteins can be used to treat any von Millebrand factor-associated
CC		disorder, regulate extracellular matrix structuring, cellular adhesion,
CC		and cell trafficking and/or migration, modulate cellular interactions,
CC		modulate cell adhesion in proliferative disorders, such as cancer,
CC		modulate the proliferation, differentiation, and/or function of cells
CC		that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC		and hematopoietic associated diseases and disorders, atelectasis,
CC		pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC		asthma and bronchiectasis, intestinal disorders, spleen associated
CC		diseases, modulate renal disorders, treat cardiovascular disorders such
CC		as ischemic heart disease, modulate the proliferation, differentiation,
CC		and/or function of bone and cartilage cells and to treat bone and/or
CC		cartilage associated diseases or disorder. They may also be used to
CC		treat disorders associated with the ovaries, and cerebral oedema,
CC		hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC		bacterial and viral meningitis, Alzheimer's disease, cerebral
CC		toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC		hydrocephalus and encephalitis, and treat hepatic disorders.
CC		note: the present sequence does not appear in the specification; it was
CC		created using information provided.
XX	SQ	Sequence 1713 BP; 505 A; 439 C; 400 G; 369 T; 0 other;
OY		Query Match 35.2%; Score 464.6; DB 21; Length 1713;
		Best Local Similarity 83.7%; Pred. No. 3.4e-66;
Matches	572; Conservative	0; Mismatch 84; Indels 27; Gaps 3.
OY	595	GCTGAGCCACTGCACGTCTGTGCACGGTCTTCATCAGCGGCACAGCCAGCGGGCGAG 654
Dd	1	GCTGAGCCACTGCACGTCTGTGCACGCCGCTCTTCATCAGCCAGCACAGCCAGGCCAGGAG 60
OY	655	GAGAGAAGATGCGCTGTCTCTGAAAGATGCTGTGGCCCTCGAGAGACACAGTGTCAGG 714
Dd	61	GAGAGAAGATGCACTGTCTCTCGAGAGATGCTGTGGCCCTCAGAGAGTAACGTTGCAGG 120
OY	715	CAGGGAATCCCAGAGCGCCAGGTTCAGCCCCCGGCTCGGGCCACCAGACCGCTGGCGCT 774
Dd	121	---TGAATGCGCGGAGCCACAGGTTCATGCCCCGCTCGGCCCTGACCGACATGCGTCTT 177
OY	775	GCGGCGCTTCGCGCCAGCGGAGCGGTTCACACGCTTCAGCCAGCCACTTATCCGTACTGTGA 834
Dd	178	GCGGCGCTTCATCCACAGCG-----AGCGATTCACAACCCACTTACCTTACTGTGA 228
OY	835	GCACAGATGCACCTGCGACCCACCATCTTCGCTTCAGACGGGAGAGAGACCCACCCCTTA 894
Dd	229	GCACAGAAATTGCGCTTCGACCCACCATCTTCATGCTGTATGGGAGAGAGACCCACCCCTTA 288
OY	895	CCAGGCGCCCTTCACACCTTCACGCTTCGGAGACCCCGACAGACAGCTGGAACCTGAACCGGGA 954
Dd	289	CCAGGCGCCCTTCACACCTTCACGCTTCAGGAGACCCCTTAGCAACAGCTGGAACCTGAACCGGGA 348

OY	955	GTMGGTGGCGGACCCCAAAACAGAACCATCTGTGACAGTGACCTGAATGATAGTCGACG	1014
Db	349	ATTCTGTGGCGCACCCCTTAACCGGACCATTCTTGACAGTACTTTATAGACAGCACCATT	408
OY	1015	GCTGGCGGCCCCCTMGCCCCCCCAGACAGTAACTCGGGGCATCAGCCCAACGTCTACGGCAG	1074
Db	409	GCTGGGGGGGCCCTTCTCCCCACGACAGTAACTCGGGGCATGACGCCCACTCTACACAG	468
OY	1075	CGGCGGCGCATMGAGSGGGCCGCGCCACTAAGCGAGGTCATCGGCCACTACCCGGG	1134
Db	469	CGGTGGGCGCATMGAGSGGGCCGCGCCCACTAAGCGAGGTCATMGGCCACTACCTGG	528
OY	1135	GTCCTCTTCCAGCACAACGACAGAGTAGGGGCCGCCCTCTGCTGTGAGGGGACCCGGCT	1194
Db	529	CTTCCTCTTCCAGCACAACGACAAAGTAAAGGGGCATCTCCCTGTAGAGGGGACCCGGCT	588
OY	1195	CCACCACACACACATCGCGGCCCTTAGAGAGGCGCAGCCATCTGGAGCAAAAGAGATAA	1254
Db	589	CCATCAGCTCGCAGATTTGCCCCACTGGA-----GAAACAAGGAGAAGAGAA	633
OY	1255	ACAGAAAGACACCCCTCTTAGG	1277
Db	634	ACAGAAAGCTACCCCTCTTAGG	656
RESULT 14			
AAA75167			
ID	AAA75167	standard; cDNA; 1713 BP.	
XX	AAA75167;		
AC	AAA75167;		
XX	15-JAN-2001	(first entry)	
DT			
DE	cdna clone encoding a murine TANGO 261 polypeptide.		
XX			
KW	TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;		
KW	cellular proliferation; cellular differentiation; cellular adhesion;		
KW	von Willebrand factor-associated disease; atelectasis; pulmonary congestion;		
KW	hematopoietic associated disease; chronic bronchitis; bronchial asthma; bronchiectasis		
KW	Oedema; emphysema; spleen associated disease; renal disorder;		
KW	intestinal disorder; ischemic heart disease; hydrocephalus;		
KW	cardiovascular disorder; inflammatory disease; inflammation; meningitis;		
KW	brain herniation; iatrogenic disease; Parkinson's disease;		
KM	Alzheimer's Disease; cerebral toxoplasmosis; encephalitis; hepatic disorder; SS.		
XX	Mus sp.		
OS			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	2..655	
FT		/tag= a	
FT		/product= "TANGO 261"	
XX			
PN	WO200052022-A1.		
XX			
PD	08-SEP-2000.		
XX			
PF	01-MAR-2000; 2000WO-US05226.		
XX			
PR	01-MAR-1999; 99US-0122458.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Barnes TM, Holtzman DA, Sharp JD, Fraser CC;		
DR	WPI; 2000-579269/54.		
XX	P-PSDB: AAB18465.		
PT	Novel human and murine secreted proteins designated TANGO 216, 261,		
XX	262, 266 and 267 useful as modulating agents of cellular processes,		
XX	e.g. for treating cancer -		



CC and/or function of bone and cartilage cells and to treat bone and/or  
CC cartilage associated diseases or disorder. They may also be used to  
CC treat disorders associated with the ovaries, and cerebral oedema,  
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
CC bacterial and viral meningitis, Alzheimer's disease, cerebral  
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
CC hydrocephalus and encephalitis, and treat hepatic disorders.  
CC note: the present sequence does not appear in the specification; it was  
CC created using information provided.

XX  
SQ Sequence 1713 BP; 506 A; 439 C; 399 G; 369 T; 0 other;

Query Match 35.2%; Score 464.6; DB 21; Length 1713;  
Best Local Similarity 83.7%; Pred. No. 3.4e-66;  
Matches 572; Conservative 0; Mismatches 84; Indels 27; Gaps 3;

QY 595 GCTGAGCCACTACAAAGCTGTCTGACAGGTCTTCATCAGCCGACCAAGCCAGGGGCGGAG 654  
Db 1 GCTGAGCCACTACAAAGCTGTCTGACAGGTCTTCATCAGCCGACCAAGCCAGGGGCGGAG 60  
QY 655 GAGGAAAGATGCTCTGTCTTCTGAGAGATGCTGTGGCCCTGCGAGAGACAGTGTCTAG 714  
Db 61 GAGGAGAGATGCTCTGTCTTCTGAGAGATGCTGTGGCCCTGCGAGAGATGCTGTCTAG 120  
QY 715 CAACGGAATCCAGAGCCGAGGCTTACGCGCCGCTGCGCCCAACGCGCCGCTGGCCGT 774  
Db 121 ---TGGAATGCCGAGACCAAGCTGTATGCCCGCTGCGCCCACTGACGAGCTGTGCTGT 177  
QY 775 GCGGCCCCCTTGCGGCGAGCGGGAGCGCTTCCACCGCTTCCAGCCCACTATCCGTACTGCA 834  
Db 178 GCGGCCCCCTTATCCAGGGG-----AGCGATTCCAACCCCACTACCCCTACTGCA 228  
QY 835 GCAGGAATGACCTGCGCACCACCATCTCGCTGTACAGCGGGGAGAGAGCCCCCACTTA 894  
Db 229 GCAGGAATGCTCTGCGCACCACCATCTCGCTGTATGGGAGGAGGCCCACTTA 288  
QY 895 CCAGGGCCCCCTGACCCCTCCAGCTTTCGGAGCCCGAGCAGAGCTGGAATGAAACCGGGA 954  
Db 289 CCAGGGCCCCCTGACCCCTCCAGCTTTCGGAGCCCGAGCAGAGCTGGAATGAAACCGGGA 348  
QY 955 GTCGGTGCAGCCACCCCAACAGAACCATCTTGCACAGTACCTGATGATAGTGCAG 1014  
Db 349 ATCTGTGCGGCGACCCCTTACCGGACATCTTGACAGTGACCTTATAGACAGACCAT 408  
QY 1015 GCTGGGGGGCCCCCTGCCCCCAGCAGTAATCGGGCATTCAGCGCCACAGTCTACGGCAG 1074  
Db 409 GCTGGGGGGCCCCCTGCCCCCAGCAGTAATCGGGCATTCAGCGCCACAGTCTACAGCAG 468  
QY 1075 CGGCGGGCGCATGAGAGGGGCGCGCCCACTACAGGAGGTGATCGGSCACTACCCGGG 1134  
Db 469 CGGTGGGCGCATGAGAGGGGCGCGCCCACTACAGGAGGTGATCGGSCACTACCCCTGG 528  
QY 1135 GTTCCTCTTCAGACAGCAGCAGTGGGCGCCCTCTTGTGAGGGGACCCGGCT 1194  
Db 529 CTCTCTCTTCAGACAGCAGCAAAAGTAAGGGGCATCTCCCTGTAGAGGGGACCCGGCT 588  
QY 1195 CCACCACACAGCATTCGGGCCCCCTTAGAGAGCGCAGCCATCTGAGCAAAAGAGATAA 1254  
Db 589 CCATCACTCGCACATTCGCCCACTGAG-----GAACAAGAGAAAGAGANA 633  
QY 1255 ACAGAAAGACACCTCTCTAGG 1277  
Db 634 ACAGAAAGTACCCCTCTAGG 656

Search completed: March 17, 2003, 12:28:53  
Job time : 370.612 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 12:18:54 ; Search time 63.5678 Seconds  
(without alignments)  
6373.035 Million cell updates/sec

Title: US-09-934-249-1

Perfect score: 1321

Sequence: 1 cgaccgcgcgtccgcgagcga.....ctgcgtagtgtgaagcag 1321

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Optical number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

Issued\_Patents\_NA: \*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCPUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	352.2	26.7	921	4	US-09-091-952A-7
2	352.2	26.7	8065	4	US-09-091-952A-6
3	284.2	21.5	867	4	US-09-091-952A-8
4	68	5.1	44377	2	US-08-804-227C-7
5	68	5.1	44377	2	US-08-804-198-1
6	67.6	5.1	4257	2	US-08-690-473-1
7	67.6	5.1	4257	4	US-09-259-821A-1
8	67.6	5.1	4257	4	US-08-843-659-1
9	67.6	5.1	12001	1	US-08-458-568A-11
10	66	5.0	152331	3	US-09-128-155-16
11	65.4	5.0	4403765	4	US-09-103-840A-2
12	64	4.8	4524	2	US-08-845-998-7
13	64	4.8	4524	3	US-09-206-537-7
14	64	4.8	4524	4	US-09-430-854-7
15	63.2	4.8	2991	3	US-08-795-430-48
16	63.2	4.8	2991	4	US-09-355-700-48
17	62.2	4.7	319	4	US-09-165-264-8
18	62.2	4.7	8438	1	US-07-945-283-1
19	62	4.7	8438	1	US-07-945-283-1
20	61.4	4.6	4060	1	US-08-308-949A-1
21	61	4.6	4411529	4	US-09-103-840A-1
22	60.4	4.6	320	4	US-09-165-264-13
23	59.8	4.5	320	4	US-09-165-264-14
24	59.4	4.5	8147	4	US-09-514-247A-9
25	59.2	4.5	320	4	US-09-165-264-7
26	59.2	4.5	71989	4	US-09-443-501A-2
27	59	4.5	801	2	US-08-770-379-16

28	59	4.5	801	4	US-08-757-669A-16	Sequence 16, Appl
29	59	4.5	801	4	US-09-298-568-3	Sequence 3, Appl
30	59	4.5	801	4	US-09-230-371A-16	Sequence 16, Appl
31	58.4	4.4	320	4	US-09-165-264-11	Sequence 11, Appl
32	58.2	4.4	15872	4	US-09-105-537-1	Sequence 1, Appl
33	58.2	4.4	43280	2	US-08-804-227C-1	Sequence 1, Appl
34	58	4.4	44377	2	US-08-804-227C-7	Sequence 7, Appl
35	58	4.4	44377	2	US-08-804-198-1	Sequence 1, Appl
36	57.6	4.4	318	4	US-09-165-264-12	Sequence 12, Appl
37	56.8	4.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl
38	56.4	4.3	1877	4	US-09-780-173A-10	Sequence 10, Appl
39	56.2	4.3	1209	6	5352575-4	Patent No. 5352575
40	56.2	4.3	1213	4	US-09-232-468A-7	Sequence 7, Appl
41	56	4.2	4488	4	US-08-406-030A-3	Sequence 3, Appl
42	56	4.2	5434	2	US-08-841-349-1	Sequence 1, Appl
43	56	4.2	12001	1	US-08-458-568A-11	Sequence 11, Appl
44	55.6	4.2	2483	1	US-08-464-340A-3	Sequence 3, Appl
45	55.6	4.2	2483	5	PCT-US94-08449A-3	Sequence 3, Appl

#### ALIGNMENTS

RESULT 1  
US-09-091-952A-7  
Sequence 7, Application US/09091952A  
Patent No. 6458532  
GENERAL INFORMATION:  
APPLICANT: Detera-Wadleigh, Sevilla D.  
Gershon, Elliot S.  
Badner, Judith A.  
Goldin, Lynn R.  
Berrettini, Wade H.  
Yoshikawa, Takeo  
Sanders, Alan R.  
Esterling, Lisa E.  
TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/091,952A  
FILING DATE: 19-Apr-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,278  
FILING DATE: 28-Oct-1996  
APPLICATION NUMBER: PCT/US97/19381  
FILING DATE: 28-Oct-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Timothy L.  
REGISTRATION NUMBER: 35,367  
REFERENCE/DOCKET NUMBER: 015280-29710005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 921 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single





```

Db 461 CTGTGGCTTTCAGACAGCGCCGACCGCGCGTGG-----CGCCTCGAGATCATGCAT 514
Qy 746 CCGCGTGGGCGCCACGACCGCTGGCGGTGCGCCCTTCCGCGGAGGAGCGTTCCAC 805
Db 515 GCCCGGGGTCAGGACAGGTTTACAGCGCGCTTCATTCAGAGGATCGCTTACG 574
Qy 806 CGCTTCCAGCCCACTATCCGTACCTCAGACAGATGCACTGCACACCCACCATCTCG 865
Db 575 CGCTTCCAGCCCACTATCCGTATGTGACAGCAGATGATCTTCTCCACCATCTCC 634
Qy 866 CTGTACAGCGGGGAGAGCCCGCCACCTACCGAGGCGCTCGACCTTCAGCTTCGGAC 925
Db 635 CTGTCCAGCGGTGAAGAGCCACCTCTTACAGGGGCGCTCGACCTTCAGCTTCGGAC 694
Qy 926 CCGAGAGAGAGCTGGAACCTGAACCGGAGTGGTGGCGGACCCCGCAACAGAACATC 985
Db 695 CCGAAGAGAGAGTGAACCTGAACCGAGAGTCCGTAGGGGCGCCCAACCGAACATA 754
Qy 986 TTGCACAGTACCTGATGATAGTCCAGG---CTGGGCGGCGCCCTCGCCCGCAGAGT 1042
Db 755 TTGTACAGTATTTAATAGACATTTGATATAGCGGGGTCCATGCCACCGACAGC 814
Qy 1043 AACTGGGCAATCAGCGCCACGCTGTACGCAAGCGGGGCGCATGAGAGGCGCGCGCC 1102
Db 815 AACTGGGCAATCAGTGAAGCAGCCTGCAGCAGTAACGGGAGATGAGAGGCGCACCCCC 874
Qy 1103 ACCTACAGCGAGTCACTGCGCACTACCGGGGTCTCTTCGACGACAGCAGACAGCAG 1161
Db 875 ACATACAGCGAGTGAATGGCGCACCGCCGCTCTTCTTCATCACACGACGAG 933

```

# RESULT 3

```

US-09-091-952A-8
; Sequence 8, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Detera-Wadleigh, Sevilla D.
; Gershon, Elliott S.
; Badner, Judith A.
; Goldin, Lynn R.
; Berretini, Wade H.
; Yoshikawa, Takeo
; Sanders, Alan R.
; Esterling, Lisa E.
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic
; Tests for Manic-Depressive Illness
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,952A
; FILING DATE: 19-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029, 278
; FILING DATE: 28-Oct-1996
; APPLICATION NUMBER: PCT/US97/19381
; FILING DATE: 28-Oct-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 015280-297100US

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; TELEEX: <Unknown>
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 867 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: cDNA
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..867
; OTHER INFORMATION: Clone 22 isoform 2 alternatively
; spliced coding region
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
; US-09-091-952A-8

```

```

Query Match 21.5%; Score 284.2; DB 4; Length 867;
Best Local Similarity 68.7%; Pred. No. 3e-45;
Matches 453; Conservative 0; Mismatches 143; Indels 63; Gaps 2;

```

```

Qy 506 CAGAGCATGAGATCAGGAGCTGAGTTGTTGATCATCATCATCATGTTGATG 565
Db 166 CCGGGCATCTTCAACTCGAGCTGAGATGTCGCCCAATCATCATCATCATGTTG 225
Qy 566 ATGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 625
Db 226 ACGGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 285
Qy 626 TTTCATGACCGCGGACAGCGAGGCGGAGAGAGAGATGCTGCTCGAAGATGC 685
Db 286 TTTCATGACCGCGGACAGCGAGGCGGAGAGAGATGCTGCTCGAAGATGC 329
Qy 686 CTGTGGCCCTTGAGAGACAGATGTCAAGGCAACGGAATCCAGAGCGGAGTCA 745
Db 330 -----GCGCAGATCATGCAT 345
Qy 746 CCGCTGGGCGCCACCGACCGCTGGCGCTGCGCCCTTCCGCGGAGGAGCGTTCC 805
Db 346 GCCCGCGGCTCAGGAGAGGTTCAAGCGCGCTTCATTCATTCAGAGGATCGCT 405
Qy 806 CGCTTCCAGCGCCACCTATCCGTACCTGACGACGAGATCGACCTGCACCATCTG 865
Db 406 CGCTTCCAGCGCCACCTATCCGTATGTGACGACGAGATGATCTTCCACCATCT 465
Qy 866 CTGTACAGCGGGAGAGAGCCCGCCACCTACCGAGGCGCTGACCTTCAGCTTCG 925
Db 466 CTGTCCAGCGGTGAAGAGCCACCTCTTACGAGGGGCGCTGACCTTCAGCTTCG 525
Qy 926 CCGAGAGCAGCTGGAATCTGAACCGGAGTGGTGGCGGACCCCGCAACAGAACATC 985
Db 526 CCGAAGAGAGAGTGAACCTGAACCGAGAGTCCGTAGGGGCGCCCAACCGAACATA 585
Qy 986 TTGCACAGTACCTGATGATAGTCCAGG---CTGGGCGGCGCCCTCGCCCGCAGAGT 1042
Db 586 TTGCACAGTATTTAATAGACATTTGATATAGCGGGGTCCATGCCACCGACGAC 645
Qy 1043 AACTGGGCAATCAGCGCCACGCTGTACGCAAGCGGGGCGCATGAGGAGCGCGCC 1102
Db 646 AACTGGGCAATCAGTGAAGAGCAGTGAAGCGGAGATGAGAGGCGCGCACCCCG 705
Qy 1103 ACCTACAGCGAGTCACTGCGCACTACCGGGGTCTCTTCACGACACGACAGCAG 1161
Db 706 ACATACAGCGAGTGAATGGCGCACCGCCGCTCTTCTTCATCACACGACGAG 764

```

```

RESULT 4
US-08-804-227C-7/c
; Sequence 7, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:

```

APPLICANT: Dehoff, Bradley S.  
APPLICANT: Kustoss, Stuart A.  
APPLICANT: Rostock, Paul R., Jr.  
APPLICANT: Sutton, Kimberly L.  
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THOMAS G. PLANT 1501  
STREET: LILLY CORPORATE CENTER  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: ASCII(DOS) Text only  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804,227C  
FILING DATE: February 21, 1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Plant, Thomas, G.  
REGISTRATION NUMBER: 35,784  
REFERENCE/DOCKET NUMBER: X-8231  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-2459  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4437 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 350..14002  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 14046..20036  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 20110..31284  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31329..36071  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 36155..41830  
US-08-804-227C-7

Query Match 5.18; Score 68; DB 2; Length 44377;  
Best Local Similarity 49.2%; Pred. No. 0.00018;  
Matches 234; Conservative 0; Mismatches 240; Indels 2; Gaps 2;

QY 5 CGCGCTCTCGAGGAGAAACCGATCTCTTGAGACTTGATGAGAGAGAGCGCGCGC 64  
DB 19499 CTGCGGTGCGGAACTGTCCTCGGACGCCGCGGCGCAGCGTCCCGCGCGGTGCGCGCGCGC 19440  
QY 65 GCGCGCGCGCGCGAGCGCTCGGTGCGGGAACCTAGCGCGCAGAGCGCTCAGCCCCGCGC 124  
DB 19439 GCGGCGAGGCGGACGCGGACGCTCTGTCCTCGCGCGCGCGCGCGCGCGCGCGCT 19380  
QY 125 GCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 184  
DB 19379 GCGGCGCGCTCGC 19320  
QY 185 CCGCGCGCGCTCGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 244  
DB 19319 GCGCGCGCGCGCGCGAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 19261  
QY 245 ECCCGCGCGCGCGCGGAACTTGGCGGCGAGCCGAGCGCGCGCGCGCGCGCGCTCC 304

DB 19260 CCGGTCGAGCAGAGTTCAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19201  
QY 305 CCGCGCGCGCGCGCGCTCTGATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 364  
DB 19200 CCGCGCGCGCGCGCGCTCGGTCCGTCAGTCGCGTGGCGAGACCCCGCGAGCACAGCCCC 19141  
QY 365 GC 424  
DB 19140 AGCGGAGCGACAGCGCGGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19082  
QY 425 ATGGGCGTCAACAGCACCG 480  
DB 19081 AAGCGGTGGCGAG 19026

RESULT 5  
US-08-804-198-1/c  
Sequence 1, Application US/08804198  
Patent No. 5945320  
GENERAL INFORMATION:  
APPLICANT: Burgett, Stanley G.  
APPLICANT: Kustoss, Stuart A.  
APPLICANT: Rao, Nagaraja R.  
APPLICANT: Richardson, Mark A.  
APPLICANT: Rostock, Paul R., Jr.  
TITLE OF INVENTION: PLATENOIDE SYNTHASE GENE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PAUL R. CANTRELL 1138  
STREET: LILLY CORPORATE CENTER  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh 7.0  
SOFTWARE: Microsoft Word 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804,198  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CANTRELL, PAUL R.  
REGISTRATION NUMBER: 36,470  
REFERENCE/DOCKET NUMBER: P9113  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3885  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4437 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 350..14002  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 14046..20036  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 20110..31284  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31329..36071  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 36155..41830  
US-08-804-198-1

Query Match 5.1%; Score 68; DB 2; Length 44377;  
Best Local Similarity 49.2%; Pred. No. 0.00018;  
Matches 234; Conservative 0; Mismatches 240; Indels 2; Gaps 2;

QY 5 GCGGCTCGAGAGGAAACCGATCTCTTGTGATGAGAGAGAGAGCGCGC 64  
DB 19449 CTCGGGTCCGAGAGCTCCCGAGCGCGCGCCAGCGTCCCGGGGGGTGGCGGGG 19440  
QY 65 GCGCGCGCGCGAGAGCGCTCGGTGGGAAAGCTACCGCAGAGCTCAGCCCCGCG 124  
DB 19439 GGGGAGAGAGGAGAGAGCGCTCTCCCGGTGCGCGCGCGCGCGCGCGGTCT 19380  
QY 125 GCACGCGCGCGCGCGCTCCAGCCATTTCCGGAGCGCCACCGGGGACATGCGGAGCG 184  
DB 19379 GCGGCGCTCCCGCTGACCGCGCGCGAGTCAAGAGAGCTCCGGAGAGAGGTGGCAC 19320  
QY 185 CCGCGGAGCTCGAGAGGAGCGCGGGGCGAGCGAGCGCGGTCCCGCGACTGAG 244  
DB 19319 GCGCGCGCGCGAGGTGAGTCCGAGCGCGAGAGCGCGCGCGGTGCGCGGT -CAGCG 19261  
QY 245 CCGCGCGCGCGCGCGGAACTTGGCGCGAGCCGAGCCGCGAGCGCGCGCGCTCC 304  
DB 19260 CCGGTCGAGAGGTCTCAGTCCGCGCGCGGTGAGCGCGGGGAGAGCCGAGCGCTGA 19201  
QY 305 CCGCGCGCGCGCTCTGATGCGGGGCGCGAGCTCGGGGCGCGCGCGCGCGCGCG 364  
DB 19200 CCGCGCGCGCTCTGCTCTGCTCACTCGGTGCGAGAGAGCCAGCGAGCGAGCGCG 19141  
QY 365 GCTCATGCGCGCTTG 424  
DB 19140 AGCGGAGGAGAGCGCGGGGTGCGCGGGGCGAGCGCGCGCGCGCGCGCGCTG -CATG 19082  
QY 425 ATGGGGGTCAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
DB 19081 AAGCGGTGCGCGCGCGCTACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGAGAC 19026

## RESULT 6

US-08-690-473-1/c  
; Sequence 1, Application US/08690473  
; Patent No. 5876923  
; GENERAL INFORMATION:  
; APPLICANT: Leopardi, Rosario  
; APPLICANT: Roizman, Bernard  
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN  
; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/690,473  
; FILING DATE: 26-JUL-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Highlander, Steven L.  
; REGISTRATION NUMBER: 37,642  
; REFERENCE/DOCKET NUMBER: ARCD:239  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
;

LENGTH: 4257 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-690-473-1

Query Match 5.1%; Score 67.6; DB 2; Length 4257;  
Best Local Similarity 48.8%; Pred. No. 0.00016;  
Matches 239; Conservative 0; Mismatches 249; Indels 2; Gaps 2;

QY 40 TGAATGAGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGTGGGAAAGC 99  
DB 2828 TGGCTGGGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGCGCG 2769  
QY 100 TAGGCGAGAGAGCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 159  
DB 2768 GCGCGCGAGAGCG 2709  
QY 160 GCGCACCGCGCGCGCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 219  
DB 2708 GCGTGGGCGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 2649  
QY 220 GCGAGCGCGGTCCCGCGACTGAGCC -CGCGCGCGCGCGCGGAACTTGGCGCGAGCGG 278  
DB 2648 GGGGCGCGCGGTCCGAGCCGCGCGCGCGCGCGCGCGCGCGCGCTCTTCTTCTTCTG 2589  
QY 279 AGCGCGCGAGCGCGCGCGCGCTTCCCGCGCGCGCGCTCTCTGATCGGGGCGCGAGC 338  
DB 2588 GGGCGCGCGCGCTCG 2529  
QY 339 TCGG -GGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 397  
DB 2528 GCGCGCGAGAGCG 2469  
QY 398 GCGCGCGCGCGCTCATGCGCGCTTGTGATGGGGGTCAACAGACCGCGCGCGCGCGCG 457  
DB 2468 GCGCGAGAGAGGGGCGAGCGCTGTGTTCGTAACAGACAGGTGCGCGCGCGCGCGCG 2409  
QY 458 GGGCAGCCCAATGTCTCTGCACTGCACTGCAACAGCGCGCTTTTGTTCAGAGCATGAG 517  
DB 2408 GCGGAGCTCGAGAGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 2349  
QY 518 ATCAAGGAGC 527  
DB 2348 CTCAGGCGC 2339

## RESULT 7

US-09-259-821A-1/c  
; Sequence 1, Application US/09259821A  
; Patent No. 6210926  
; GENERAL INFORMATION:  
; APPLICANT: LEOPARDI, ROSARIO  
; APPLICANT: ROIZMAN, BERNARD  
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS  
; TITLE REFERENCE: ARCD:317  
; CURRENT APPLICATION NUMBER: US/09/259,821A  
; CURRENT FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 08/690,473  
; PRIOR FILING DATE: 1996-07-26  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4257  
; TYPE: DNA  
; ORGANISM: HERPES VIRUS, TYPE 1  
US-09-259-821A-1

Query Match 5.1%; Score 67.6; DB 4; Length 4257;  
Best Local Similarity 48.8%; Pred. No. 0.00016;  
Matches 239; Conservative 0; Mismatches 249; Indels 2; Gaps 2;  
QY 40 TGAATGAGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGTGGGAAAGC 99

; PRIOR APPLICATION DATA;  
 ; APPLICATION NUMBER: US 09/065 146

```

1      RESULT 8
2      US-08-843-659-1/c
3      : Sequence 1, Application US/08843659
4      : Patent No. 6218103
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: Leopardi, Roarid
8      : APPLICANT: Roizman, Bernard
9      : TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS
10     : TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
11     : NUMBER OF SEQUENCES: 6
12     : CORRESPONDENCE ADDRESS:
13     : ADDRESSEE: Arnold, White & Durkee
14     : STREET: P.O. Box 4433
15     : CITY: Houston
16     : STATE: Texas
17     : COUNTRY: United States
18     : ZIP: 77210
19     :
20     : COMPUTER READABLE FORM:
21     : MEDIUM TYPE: Floppy disk
22     : COMPUTER: IBM PC compatible
23     : OPERATING SYSTEM: PC-DOS/MS-DOS
24     : SOFTWARE: PatentIn Release #1.0, Version #1.30
25     : CURRENT APPLICATION DATA:
26     : APPLICATION NUMBER: US/08/843,659
27     : FILING DATE: Concurrently Herewith
28     : CLASSIFICATION: 435
29     : ATTORNEY/AGENT INFORMATION:
30     : NAME: Highlander, Steven L.
31     : REGISTRATION NUMBER: 37,642
32     : REFERENCE/DOCKET NUMBER: ARS:519
33     : TELECOMMUNICATION INFORMATION:
34     : TELEPHONE: (512) 418-3000
35     : TELEFAX: (512) 474-7577
36     : INFORMATION FOR SEQ. ID NO: 1:
37     : SEQUENCE CHARACTERISTICS:
38     : LENGTH: 4257 base pairs
39     : TYPE: nucleic acid

```



OTHER INFORMATION: "\*" bases at various positions throughout the sequence  
OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match  
Best Local Similarity 5.0%; Score 65.4; DB 4; Length 4403765;  
Matches 219; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

QY 1 CGACCGCGGCTCGAGAGGAAACCGATCTCTTGGACTTGAATGAGAGAGAGCGG 60  
DB 3936911 CGCCACCCCTTCCGCGCGCCGCCCTTACCGGCGGCCCGCGGCGCCGCCG 3936852  
QY 61 CGGCGCGCGCGCGCGCGCGCGCTGCGCTGCGGAAAGCTAGCGGAGAGAGAGCGCC 120  
DB 3936851 CGGTGATATGTCGCGCTGCGCGCGCTCCCGCGGACCGCGGACACCTTATGAC 3936792  
QY 121 GCGCGAGAGCGCGCGCGCGCTGCGAGCCATTTCCGAGCCACCGCGGCGACTGCGC 180  
DB 3936791 CGGCACCGCGCGCGCGCGCGCTGCGACCTTGTCCGCGCGCGCGCGCGCGCG 3936732  
181 ACGCCCGCGCGCTCGCGAGGAGCGCGGAGCGGAGCGGAGCGGAGCGCGCGCGC 240  
DB 3936731 TCGCGCGCGCGCTCGCGGCGCGCGCGCGCGCGCGCGCTCGCTCGCGTGC 3936672  
QY 241 TGAGCCCGCGCGCGCGCGCGGAACTGGCGGAGAGCGCGCGGAGCGCGCGCG 300  
DB 3936671 CGCGCGCGCGCGCTGCGCGCTGCTGAGCTGAGAGGCTGCGCGCGCGCGCT 3936612  
QY 301 CTCGCCCGCGCGCGCTCTGCTGATGCGGCGCGCGCGCGCGCGCGCGCGCGCG 360  
DB 3936611 CGGACCGCGCTTGGCGCGCTTGGCCAGGCGGCGCGCGCGCGCGCGCGCG 3936552  
QY 361 CCG 420  
DB 3936551 CCG 3936492  
QY 421 CTGATGGGAGGCTCAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3936437  
DB 3936491 CGGACGCTGTGTGGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 3936437

## RESULT 12

US-08-845-998-7/c  
Sequence 7, Application US/08845998  
Patent No. 5879892

## GENERAL INFORMATION:

APPLICANT: Van Baren, Nicolas  
APPLICANT: Coulie, Pierre G.  
APPLICANT: De Smet, Charles  
APPLICANT: Lucas, Sophie  
APPLICANT: Boon, Thierry  
TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/845,998  
FILING DATE:

## CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Van Amsterdam, John R.  
REGISTRATION NUMBER: 40,212  
REFERENCE/DOCKET NUMBER: L0461/7008

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)720-3500  
TELEFAX: (617)720-2441  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4524 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:

NAME/KEY: CDS  
LOCATION: 174..1433  
US-08-845-998-7

Query Match  
Best Local Similarity 4.8%; Score 64; DB 2; Length 4524;  
Matches 218; Conservative 0; Mismatches 220; Indels 2; Gaps 2;

QY 166 CCGCGGCGACTGCGCGAGCGCCCGCGGCGTGCAGAGGAGCGCGGAGCGCGAG 225  
DB 897 CCACCTGCG 839  
QY 226 CCGCGTCCCGCGCGACTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 285  
DB 838 GCGACAGCG 779  
QY 286 CGAGCGCGGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 345  
DB 778 TCGCAGAGGTGCG 719  
QY 346 GC-CGCGGAGAGCG 404  
DB 718 CCG 659  
QY 405 CGCGCTCCATGACCGCTTGTATGCGGAGTCAACAGACCGCGCGCGCGCGCGCG 464  
DB 658 CCG 599  
QY 465 CCAATGCTCTCTGCAAGTGAACGCTCTTGTTCAGAGCATGAGATACCG 524  
DB 598 CCG 539  
QY 525 AGCTGAGTTTGTTCAGATCATCATGCTGTGTGTATGATGATGATGATGATGATG 584  
DB 538 GAGGAGATGTTGTCAGACATGCGCGGTTGAGCGCGCTGTGTGTGTGTGTGTGTG 479  
QY 585 TCAGTGCCTGTGAGCCAC 604  
DB 478 TGTGTGTGCGCGCGAGAGGAC 459

## RESULT 13

US-09-206-537-7/c  
Sequence 7, Application US/09206537  
Patent No. 6130052

## GENERAL INFORMATION:

APPLICANT: Van Baren, Nicolas  
APPLICANT: Coulie, Pierre G.  
APPLICANT: De Smet, Charles  
APPLICANT: Lucas, Sophie  
APPLICANT: Boon, Thierry  
TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02210

```

1 COMPUTER READABLE FORM:
2 MEDIUM TYPE: FLOPPY disk
3 COMPUTER: IBM PC compatible
4 OPERATING SYSTEM: PC-DOS/MS-DOS
5 SOFTWARE: PatentIn Release #1.0, Version #1.25
6 CURRENT APPLICATION DATA:
7 APPLICATION NUMBER: US/09/206,537
8 FILING DATE:
9 CLASSIFICATION:
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: 08/845,998
12 FILING DATE:
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Van Amsterdam, John R.
15 REGISTRATION NUMBER: 40,212
16 REFERENCE/DOCKET NUMBER: L0461/7008
17 TELECOMMUNICATION INFORMATION:
18 TELEPHONE: (617)720-3500
19 TELEFAX: (617)720-2441
20 INFORMATION FOR SEQ ID NO: 7:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 4524 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: double
25 TOPOLOGY: linear
26 MOLECULE TYPE: cDNA
27 HYPOTHETICAL: NO
28 ANTI-SENSE: NO
29 FEATURE:
30 NAME/KEY: CDS
31 LOCATION: 174..1433
32 US-09-206-537-7

```

MOLECULE TYPE: DNA (genomic)

Search completed: March 17, 2003, 14:12:34  
Job time : 1067.57 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 13:40:19 ; Search time 112.606 Seconds  
(without alignments)  
8233.708 Million cell updates/sec

Title: US-09-934-249-1

Perfect score: 1321

Sequence: 1 cgaccgcgcttcgcgagcga.....ctgcgtagtgtaaaagcag 1321

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 501302 seqs, 350932545 residues 1002604

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1321	100.0	1321	10	US-09-934-249-1
2	861	65.2	861	10	US-09-934-249-3
3	803.2	60.8	969	9	US-09-796-753-55
4	790.4	59.8	1066	9	US-10-098-841-71
5	616.4	46.7	878	10	US-09-934-249-12
6	589.2	44.6	1583	10	US-10-000-256A-32
7	493.6	37.4	693	10	US-09-934-249-14
8	466.2	35.3	1713	9	US-09-796-753-57
9	349	26.4	8093	10	US-09-934-249-16
10	341.8	25.9	475	10	US-09-934-249-15
11	182	13.8	368	10	US-09-783-590-3464
12	78	5.9	4020	9	US-09-796-679-5
13	75.4	5.7	2561	12	US-10-023-529-48
14	75.4	5.7	2561	12	US-10-023-529-48
15	75.4	5.7	2561	12	US-10-023-529-48
16	72.2	5.5	2307	9	US-09-893-519A-87
17	72.2	5.5	2561	9	US-09-976-740-48
18	72.2	5.5	2561	12	US-10-023-529-48
19	72.2	5.5	2561	12	US-10-023-529-48

c	20	68	5.1	884.1	9	US-09-954-531-1351	Sequence 1351, Ap
c	21	67.6	5.1	4257	9	US-09-825-288A-1	Sequence 1, Appl
c	22	67.4	5.1	43058	10	US-09-954-456-292	Sequence 292, App
c	23	67.4	5.1	43058	10	US-09-954-456-292	Sequence 328, App
c	24	67.4	5.1	43058	10	US-09-880-107-3950	Sequence 3950, Ap
c	25	67	5.1	12733	9	US-10-032-393-47	Sequence 47, Appl
c	26	67	5.1	12739	9	US-10-032-393-8	Sequence 8, Appl
c	27	66.6	5.0	1614	9	US-09-976-740-45	Sequence 45, Appl
c	28	66.6	5.0	1614	12	US-10-023-529-45	Sequence 45, Appl
c	29	66.6	5.0	12425	9	US-09-976-740-50	Sequence 50, Appl
c	30	66.6	5.0	12425	9	US-10-023-529-50	Sequence 50, Appl
c	31	66.6	5.0	12425	12	US-10-023-529-50	Sequence 50, Appl
c	32	66.6	5.0	12425	12	US-10-023-529-50	Sequence 50, Appl
c	33	66.4	5.0	42989	9	US-09-799-462A-17	Sequence 17, Appl
c	34	66.4	5.0	42989	9	US-10-125-767-17	Sequence 17, Appl
c	35	66.4	5.0	42989	9	US-09-836-911A-17	Sequence 17, Appl
c	36	66	5.0	15231	9	US-10-095-407-16	Sequence 16, Appl
c	37	65.8	5.0	42999	9	US-09-799-462A-17	Sequence 17, Appl
c	38	65.8	5.0	42999	9	US-10-125-767-17	Sequence 17, Appl
c	39	65.8	5.0	42999	9	US-09-836-911A-17	Sequence 17, Appl
c	40	64.4	4.9	1614	12	US-10-023-529-45	Sequence 45, Appl
c	41	64.4	4.9	1614	12	US-10-023-529-45	Sequence 45, Appl
c	42	64.4	4.9	12425	9	US-09-976-740-50	Sequence 50, Appl
c	43	64.4	4.9	12425	9	US-09-976-740-50	Sequence 50, Appl
c	44	64.4	4.9	12425	12	US-10-023-529-50	Sequence 50, Appl
c	45	64.4	4.9	12425	12	US-10-023-529-50	Sequence 50, Appl

## ALIGNMENTS

RESULT 1  
US-09-934-249-1  
Sequence 1, Application US/09934249  
Patent No. US20020115081A1  
GENERAL INFORMATION:  
APPLICANT: Lee, Richard T.  
APPLICANT: Landschultz, Katherine T.  
APPLICANT: Thurl, Thomas G.  
APPLICANT: Thompson, John F.  
APPLICANT: Kennedy, Scott P.  
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS  
FILE REFERENCE: P0738/7001/ERP/KA  
CURRENT APPLICATION NUMBER: US/09/934, 249  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: US 60/227, 159  
PRIOR FILING DATE: 2000-08-22  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 1321  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (413)...(1273)  
US-09-934-249-1  
Query Match 100.0%; Score 1321; DB 10; Length 1321;  
Best Local Similarity 100.0%; Pred. No. 2.7e+242;  
Matches 1321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGACCGCGGCTCTCGAGCGAACCAGCATCTTGACTGAATGAGAGAGAGAGCGG 60  
DB 1 CGACCGCGGCTCTCGAGCGAACCAGCATCTTGACTGAATGAGAGAGAGAGCGG 60  
QY 61 CG 120  
DB 61 CG 120  
QY 121 GGGGCGAGCG 180  
DB 121 GGGGCGAGCG 180



QY	893	TACACAGGCCCCCTGCACCTCTCAGACTTCGGAGACCCCGAGAGAGCTGTAACCTGAACCGG	952
Db	481	TACACAGGCCCCCTGCACCTCTCAGACTTCGGAGACCCCGAGAGAGCTGTAACCTGAACCGG	540
QY	953	GAGTCGGTGGCGCGACCCCAACAGAACCATTTTCGACAGTGCCTGATGGATAGTGCC	1012
Db	541	GAGTCGGTGGCGCGACCCCAACAGAACCATTTTCGACAGTGCCTGATGGATAGTGCC	600
QY	1013	AGGCTGGGCGGCCCCCTGCACCCCGACAGATTAATCTGGGCGATCAGGCCACGTCTACGAC	1072
Db	601	AGGCTGGGCGGCCCCCTGCACCCCGACAGATTAATCTGGGCGATCAGGCCACGTCTACGAC	660
QY	1073	AGCGGGGGGGCGCATGAGAGGGGGCGCGCCCACTACAGCGAGGTCACTGGGCACCTACCGG	1133
Db	661	AGCGGGGGGGCGCATGAGAGGGGGCGCGCCCACTACAGCGAGGTCACTGGGCACCTACCGG	720
QY	1133	GGGTCTCTCTTCACAGCACACAGACAGTGGGGCGCCCTCTTGTCTGAGGGGACCCGG	1193
Db	721	GGGTCTCTCTTCACAGCACACAGACAGTGGGGCGCCCTCTTGTCTGAGGGGACCCGG	780
QY	1193	CTTCACCAACACACATCTGGGGCCCTAGAGAGGGCAGCCATCTGGAGCAAAAGAAAGAT	1253
Db	781	CTTCACCAACACACATCTGGGGCCCTAGAGAGGGCAGCCATCTGGAGCAAAAGAAAGAT	840
QY	1253	AAACAGAAAGACACCTCTTC	1273
Db	841	AAACAGAAAGACACCTCTTC	861

```

1 RESULT 3
2 US-09-796-753-55
3
4 ; Sequence 55, Application US/09796753
5 ; Publication No. US20030027998A1
6 ; GENERAL INFORMATION:
7 ; APPLICANT: McCarthy, Sean A.
8 ; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
9 ; FILE REFERENCE: 7853-227-999
10
11 CURRENT APPLICATION NUMBER: US/09/796,753
12
13 CURRENT FILING DATE: 2001-03-01
14
15 PRIOR APPLICATION NUMBER: 09/183,175
16
17 PRIOR FILING DATE: 1998-10-30
18
19 PRIOR APPLICATION NUMBER: 09/223,094
20
21 PRIOR FILING DATE: 1998-12-30
22
23 PRIOR APPLICATION NUMBER: 09/223,546
24
25 PRIOR FILING DATE: 1998-12-30
26
27 PRIOR APPLICATION NUMBER: 09/224,246
28
29 PRIOR FILING DATE: 1998-12-30
30
31 PRIOR APPLICATION NUMBER: 09/259,388
32
33 PRIOR FILING DATE: 1999-02-26
34
35 PRIOR APPLICATION NUMBER: 60/122,458
36
37 PRIOR FILING DATE: 1999-03-01
38
39 PRIOR APPLICATION NUMBER: 09/312,359
40
41 PRIOR FILING DATE: 1999-05-14
42
43 PRIOR APPLICATION NUMBER: 09/336,536
44
45 PRIOR FILING DATE: 1999-06-18
46
47 PRIOR APPLICATION NUMBER: 09/344,667
48
49 PRIOR FILING DATE: 1999-06-29
50
51 PRIOR APPLICATION NUMBER: 09/345,464
52
53 PRIOR FILING DATE: 1999-06-30
54
55 PRIOR APPLICATION NUMBER: 09/365,164
56
57 PRIOR FILING DATE: 1999-07-30
58
59 PRIOR APPLICATION NUMBER: 09/399,723
60
61 PRIOR FILING DATE: 1999-09-20
62
63 PRIOR APPLICATION NUMBER: 09/409,634
64
65 PRIOR FILING DATE: 1999-09-30
66
67 PRIOR APPLICATION NUMBER: 09/471,179
68
69 PRIOR FILING DATE: 1999-12-23
70
71 PRIOR APPLICATION NUMBER: 09/474,071
72
73 PRIOR FILING DATE: 1999-12-29
74
75 PRIOR APPLICATION NUMBER: 09/474,072
76
77 PRIOR FILING DATE: 1999-12-29
78
79 PRIOR APPLICATION NUMBER: 09/514,010
80

```

```

? PRIOR FILLING DATE: 2000-02-25
? PRIOR APPLICATION NUMBER: 09/516,745
? PRIOR FILLING DATE: 2000-03-01
? PRIOR APPLICATION NUMBER: 09/572,002
? PRIOR FILLING DATE: 2000-05-14
? PRIOR APPLICATION NUMBER: 09/597,993
? PRIOR FILLING DATE: 2000-06-19
? PRIOR APPLICATION NUMBER: 09/599,596
? PRIOR FILLING DATE: 2000-06-22
? PRIOR APPLICATION NUMBER: 09/630,334
? PRIOR FILLING DATE: 2000-07-31
? PRIOR APPLICATION NUMBER: 09/606,565
? PRIOR FILLING DATE: 2000-06-29
? PRIOR APPLICATION NUMBER: 09/606,317
? PRIOR FILLING DATE: 2000-06-29
? PRIOR APPLICATION NUMBER: 09/665,666
? PRIOR FILLING DATE: 2000-09-20
? PRIOR APPLICATION NUMBER: 09/677,751
? PRIOR FILLING DATE: 2000-09-30
? NUMBER OF SEQ ID NOS: 162
? SEQ ID NO 55
? LENGTH: 969
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (6)...(761)
US-09-796-753--55

```

[illegible]



LOCATION: (20)...(841)  
US-09-934-249-12

[illegible][illegible]

```

RESULT 6
US-10-000-256A-32
; Sequence 32, Application US/10000256A
; Publication No. US20030039983A1
; GENERAL INFORMATION:

```

Query Match	44.6%;	Score 589.2;	DB 9;	Length 1583;
Best Local Similarity	99.5%;	Pred. No. 1.9e-103;		
Matches 591;	Conservative	0;	Mismatches 3;	Indels 0;
			Gaps	0

OY	728	GAGCGGAGAGTCTACGGCCCGCCCTCGAGCCACAGACGGCTGAGCGTGCAGCCCTTTCGCG	787
Db	825	GGCGCCAGAGTCTACGGCCCGCCCTCGAGCCACAGACGGCTGAGCGTGCAGCCCTTTCGCG	884
OY	788	CAGCGGAGAGGCGCTTCACCGCGCTTCAGAGCCACCATCTACCTGACAGCAGATCGAC	847
Db	885	CAGCGGAGAGGCGCTTCACCGCGCTTCAGAGCCACCATCTACCTGACAGCAGATCGAC	944
OY	848	CTGCGCACCCACACATCTCGCTGTGTACAGACGGGAGAGAGCCCCACCTACAGAGGCCCTTGC	907
Db	945	CTGCGCACCCACACATCTCGCTGTGTACAGACGGGAGAGAGCCCCACCTACAGAGGCCCTTGC	1004
OY	908	ACCCCTCAGCTTCGGGAGCCCGCGAGCAGACAGTGAACCTAAACCGGAGAGTGGTGGCGCA	967
Db	1005	ACCCCTCAGCTTCGGGAGCCCGCGAGCAGACAGTGAACCTAAACCGGAGAGTGGTGGCGCA	1064
OY	968	CCCCCAACAGAACCATCTTCGACAGTGAACCTGATGATAGTGCAGAGCTGGGCGGCCCC	1027
Db	1065	CCCCCAACAGAACCATCTTCGACAGTGAACCTGATGATAGTGCAGAGCTGGGCGGCCCC	1124
OY	1028	TGCCCCCCACAGCAGTAACTGTGGGCAATCAGCGCCACGTGCTAACGCGAGCGGCGGCGATG	1087
Db	1125	TGCCCCCCACAGCAGTAACTGTGGGCAATCAGCGCCACGTGCTAACGCGAGCGGCGGCGATG	1184
OY	1088	GAGGGGGCGCGCCACCTCAGCAGCAGAGGTCATGGGCACTACCCGGGGTCTCTTTCAG	1147
Db	1185	GAGGGGGCGCGCCACCTCAGCAGCAGAGGTCATGGGCACTACCCGGGGTCTCTTTCAG	1244
OY	1148	CACGACGAGAGCAGTGGGCGCGCCCTCTTCTTGAAGGGAGCCCGGCTCCACACACACAC	1207
Db	1245	CACGACGAGAGCAGTGGGCGCGCCCTCTTCTTGAAGGGAGCCCGGCTCCACACACACAC	1304
OY	1208	ATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAAAGAGAGATTAACAGAAAGGACAC	1267
Db	1305	ATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAAAGAGAGATTAACAGAAAGGACAC	1364
OY	1268	CCTCTCAGAGGTTCGCCAGGGGGGGCGGGGCTGGGGCTGCTCAGTGA AAAAGCGC	1321
Db	1365	CCTCTCAGAGGTTCGCCAGGGGGGGCGGGGCTGGGGCTGCTCAGTGA AAAAGCGC	1418

RESULT 7  
US-09-934-249-14/c  
; Sequence 14, Application US/09934249  
; Patent No. US20020115081A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Richard T.  
; APPLICANT: Landschulz, Katherine T  
; APPLICANT: Turi, Thomas G.  
; APPLICANT: Thompson, John F.  
; APPLICANT: Kennedy, Scott P.

QY	776	CCGGCCCTTCGCCCAGCGGGAGCGCTCCACCGCTTCACGGCCCAACCTATCGTACCTCGAG	835
Db	693	CGCGCTTTCGCCCAAGCGGGAGCGGTTCCACGCTTCACGGCCCAACCTATCGTACCTCGAG	634
QY	836	CACGAGATCGACCTGCCACCCACCATTCCGCTGCACAGCGGGAGAGCCCCACCCCTAC	895
Db	633	CACGAGATCGA-CTCCGCGCCCAACCATCTCGCTGTGCAGACGGG--AGAGGCCCCACCCCTAC	577
QY	896	CAGGCGCCCTTCGACCCCTCCAGCTTCGGGAGCCCGAGCAGCAGCTGGAACCTGAAACCGGGAG	955
Db	576	CAGGCGCCCTTCGACCCCTCCAGCTTCGGGAGCCCGAGCAGCAGCTGGAACCTGAAACCGGGAG	517
QY	956	TCGGTCCGCGCACCCCCCAACGAACCACTTCCTGCAGCTGACCTGATGATAGTGCAGG	1015
Db	516	TCGGTCCGCGCACCCCCCAACGAACCACTTCCTGCAGCTGACCTGATGATAGTGCAGG	457
QY	1016	CTGGGGGGGCCCCGCCCCCAGCAGTAACTCGGGCATCGAGCGCCAGCTGCTACGGCAGC	1075
Db	456	CTGGGGGGGCCCCGCCCCCAGCAGTAACTCGGGCATCGAGCGCCAGCTGCTACGGCAGC	397
QY	1076	GCGGGGGCGCATGGAGGGGCGCGCCGCTACACTACAGCGAGTATGGGCACTACCCGGGG	1135
Db	396	GCGGGGGCGCATGGAGGGGCGCGCCGCTACACTACAGCGAGTATGGGCACTACCCGGGG	337
QY	1136	TCTCTCTCTCAGACGACGAGCAGTGGGCGCGCCCTCTTCTCTGGA-GGGGACCGGGCT	1194
Db	336	TCTCTCTCTCAGACGACGAGCAGTGGGCGCGCCCTCTTCTCTGGAAGGGGACCGGGTT	277
QY	1195	CCACCCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAAAGATTA	125
Db	276	CCCCCAACACACATTCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAAAGATTA	217
QY	1255	ACAGAAAGAGACACCTCTCTTAGGCTCCCAAGGGGGGGCGGGGCTGCGTAGGTGAA	131
Db	216	ACAGAAAGAGACACCTCTCTTAGGCTCCCAAGGGGGGGCGGGGCTGCGTAGGTGAA	157
QY	1315	AAGCGAG-1321	
Db	156	AAGCGAG-150	

RESULT 8  
US-09-796-753-57  
Sequence 57, Application US/09796753  
Publication No. US20030027998A1  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
FILE REFERENCE: 7853-227-899  
CURRENT APPLICATION NUMBER: US/09/756,753

```

CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 09/183,175
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 09/223,094
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/223,546
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/224,246
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/259,388
PRIOR FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: 60/122,458
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: 09/312,359
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/336,536
PRIOR FILING DATE: 1999-06-18
PRIOR APPLICATION NUMBER: 09/342,687
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: 09/345,464
PRIOR FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: 09/365,164
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/399,723
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 09/409,634
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 09/471,179
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: 09/474,071
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 09/474,072
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 09/514,010
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: 09/516,745
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 09/572,002
PRIOR FILING DATE: 2000-05-14
PRIOR APPLICATION NUMBER: 09/597,993
PRIOR FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: 09/599,566
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 09/630,334
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: 09/606,565
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: 09/606,317
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: 09/665,666
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 09/677,751
PRIOR FILING DATE: 2000-09-30
NUMBER OF SEQ ID NOS: 162
SEO ID NO 57
LENGTH: 1713
TYPE: DNA
ORGANISM: Mouse
FEATURE:
NAME/KEY: CDS
LOCATION: (2)...(652)
US-09-796-753-57

Query Match      35.3%; Score 466.2; DB 9; Length 1713;
Best Local Similarity 83.9%; Pred. No. 4,1e-80;
Matches 573; Conservative 0; Mismatches 83; Indels 27; Gaps 3

```

```

Db 61 GAGAGAGATGACTGTCCTGCGAAGATGCTTGCGCCCTCAGAGATACGATGTCAG 120
QY 715 CAAGGAATCCCAAGCGGAGGTCTACGCCCGCTCGGCCCAAGCCGCTGCGCT 774
Db 121 ---TGGAATCCCGAGCGACAGGTCTATGCCCGCTCGGCCCACTACCGACTCGTGT 177
QY 775 GCCGCTTGGCCCGGAGGAGCGCTTCACCGCTTCAGCGCCACTATCCCTACGCA 834
Db 178 GCGCCCTTATCCAGGGG-----AGCCGATTCCAAACCCACTTACCTTACTGCA 228
QY 835 GCACGAATGACCTGCGACACCCACCATCTGCTGTACAGCGGGAGAGAGCCCGCACTA 894
Db 229 GCAGGAATGCTGCTGCGACACCATCTGCTGTATGATGGAGAGAGCCCGCACTA 288
QY 895 CCAGGCGCTTGCACCTCTCCAGCTTCGGGAGCCCGAGCAGCAGTGGAACTGAACCGGGA 954
Db 289 CCAGGCGCTTGCACCTCTCCAGCTTCGGGAGCCCGAGCAGCAGTGGAACTGAACCGGGA 348
QY 955 GTCGGTGGCGGACACCCCAACAGAACATCTTGACAGTGAACCTGATGATGTCGAG 1014
b 349 ATCTGTGCGCGGACACCCCTTAACCGGACATCTTGACAGTGAACCTGATGATGTCGAG 408
QY 1015 GCTGGGCGCCCTGCGCCCGCCAGCAGTAACTGCGGATCAGCGCCAGCTGTCAGGCGAG 1074
Db 409 GCTGGGCGCCCTGCTGCGCCCGCCAGCAGTAACTGCGGATCAGCGCCAGCTGTCAGGCGAG 468
QY 1075 CGGGGGCGCATGAGAGGGGCGCGCCCGCTTACAGGAGGTCTATCGGCCACTACCGGG 1134
Db 469 CGGGGGCGCATGAGAGGGGCGCGCCCGCTTACAGGAGGTCTATCGGCCACTACCGGG 528
QY 1135 GTCTCTCTTCAGACAGCAGCAGCAGTGGGCGCGCCCTCTTGTGAGAGGAGCCGCGCT 1194
Db 529 CTCTCTCTTCAGACAGCAGCAGCAGTGGGCGCGCCCTCTTGTGAGAGGAGCCGCGCT 588
QY 1195 CCACCAACACACATCGCGCCCTTACAGAGCGAGCAGCATCTGAGCAAGAAAGATTA 1254
Db 589 CCACTACTCGCAGATGCGCCCACTGGA-----GAACAAGAGAAAGAGAA 633
QY 1255 ACAGAAGGACACCTCTCTAGG 1277
Db 634 ACAGAAGGTCACCCCTCTAGG 656

```

```

RESULT 9
US-09-934-249-16
; Sequence 16, Application US/09934249
; Patent No. US20020115081A1
GENERAL INFORMATION:
APPLICANT: Lee, Richard T.
APPLICANT: Landschultz, Katherine T.
APPLICANT: Turi, Thomas G.
APPLICANT: Thompson, John F.
APPLICANT: Kennedy, Scott P.
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
FILE REFERENCE: P0738/7001/ERP/KA
CURRENT APPLICATION NUMBER: US/09/934, 249
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 60/227,159
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 16
LENGTH: 8093
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (6477)..(6477)
OTHER INFORMATION: c or t/u
NAME/KEY: unsure
LOCATION: (6837)..(6837)
OTHER INFORMATION: a or c

```

US-09-934-249-16

Query Match 26.48; Score 349; DB 10; Length 8093;  
 Best Local Similarity 73.78; Pred. No. 8.2e-58;  
 Matches 474; Conservative 0; Mismatches 160; Indels 9; Gaps 2;

```

QY 522 CGGAGCTGGAATTTGTTACAGATCATCATCGTGTGTATGATGATGATGATGATG 581
Db 250 CGGAGCTGGAATTTGTTACAGATCATCATCGTGTGTATGATGATGATGATGATG 309
QY 582 TGATCAGTGTGCTGCTGAGCCACTACAAAGCTGTGACAGGCTCTTCAATCAGCCGCA 641
Db 310 TCATCTGTGCTGCTGCTGAGCCACTACAAAGCTGTGACAGGCTCTTCAATCAGCCGCA 369
QY 642 GCCAGGGGCGGAGAGAGAGATGCTCTGTCTTCAAGAGATGCTGTGCTTGGCTTCGAGA 701
Db 370 ACCAGAGCCGAGGAGGAGAGAGAGGCTGCGCAGAGAGAGGCTGCTGTGCTTCAAGACA 429
QY 702 GCACAGTGTGAGGCAAGGAAATCCAGAGCCGAGTCTACGCCCGCTGCGCCCAACCG 761
Db 430 GCGCGGACCGCGGCTGAG-----CGCTTGGAGATCATCATGCCCCGCGGTCAGAGG 483
QY 762 ACCGCTGCGCGCTGCGCCCTTTCGCCAGCGGAGAGCGCTTTCACCGCTTCAGGCCACT 821
Db 484 ACAGGTTACAGCGCGCGCTTTCAGGAGAGGATGCTTTCAGCGCGCTTTCAGGCCACT 543
QY 822 ATCGCTACTGACAGAGATGACATGCTGCGCAGCCAGCATCTGCTGTCAAGCGGAGAG 881
Db 544 ACCCTTATGTCACAGAGATGATATCTTCTCCACCATCTCTCTGCTGCGCAGGCTGAG 603
QY 882 AGCCCGCACCTTACAGAGGCGCCCTGACACCTTCAGCCTTGGAGAGCGGAGAGCTGG 941
Db 604 AGCCACTCTTACAGAGGCGCCCTGACACCTTCAAGCTTCCGAGACCTTACAGAGAGATG 663
QY 942 AACTGAACCGGAGATGCTGCGCGCACCCCAACAGAAACCATCTTTCAGAGTGA 1001
Db 664 AACTGAACCGGAGATGCTGCGCGCACCCCAACAGAAACCATCTTTCAGAGTGA 723
QY 1002 TGGATATGTCAGG---CTGGGCGCGCGCTGCGCCCGCCAGCAGTATCTGCGGATCAGCG 1058
Db 724 TAGACATTTGTTATATAGCGGGGCTCATGCTCCACAGAGAGAACTCGGGATCACTG 783
QY 1059 CCAGCTGCTACGAGCGGCGCGCATGAGAGGCGCGCCCGCACTTACAGAGAGTCA 1118
Db 784 CAAGCAGCTGAGCAGTACGAGGAGATGAGAGGCGCGCCCGCACTTACAGAGAGTGA 843
QY 1119 TGGGCACTACCGGGGCTCTCTTCCAGCAGCAGCAGAGCAG 1161
Db 844 TGGGCACTACCGAGGCGCTCTTCTCTCATACACAGCGCAG 886

```

```

RESULT 10
US-09-934-249-15
; Sequence 15, Application US/09934249
; Patent No. US20020115081A1
GENERAL INFORMATION:
APPLICANT: Lee, Richard T.
APPLICANT: Landschultz, Katherine T.
APPLICANT: Turi, Thomas G.
APPLICANT: Thompson, John F.
APPLICANT: Kennedy, Scott P.
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
FILE REFERENCE: P0738/7001/ERP/KA
CURRENT APPLICATION NUMBER: US/09/934, 249
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 60/227,159
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 475
TYPE: DNA

```

```

RESULT 11
US-09-783-590-3464
; Sequence 3464, Application US/09783590
; Patent No. US20020110850A1
;
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2.2
FILE REFERENCE: PO-16.2C1
CURRENT APPLICATION NUMBER: US/09/783,590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3464
LENGTH: 368
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (103)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (225)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (279)

```

45 GAGGAGGAGCGCGCGCGCGCGCGGAGGCGCTCGGCTGGGGAAGCTAGCG 104





```
QY 176 TCCGACGCCCCCGGGGCTGCCAGGAGAGCCGGGGGGGGCGCAGCGGAGCCGGCTCCG 235
DB 755 GCGCGGCTCGGGGGCGGGGCGAGGGGGCGCTGTGGCG-CCAGGGGGGGCGGGCGG 697
QY 236 CCGACTGAGCCCCCGGGCGCCCGGAACTTGGCGCGAGCCCGAGCCCGCGGGG 295
DB 696 CCGGGGGCGCCCGGGGCGGGGGCGGGCGAGCGGGCGCGGGCGGGCGGGGG 637
QY 296 CCGGCTTCCCGCGGGCGGCTCTGCATGCGGGGCCCACTCCGGGGCGGGCGGA 355
DB 636 CCGGGGTGGGCGGGCGGGGGCGGGCGGCG-GGCGGGGGCCCCCGGGGGGGCGGGCGGGG 578
QY 356 GCGCCCCCGGGCGCGCGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCTCATG 415
DB 577 GCGGGCGGGGGTGGCTCCGGCGGGGGCGGCTGGACGGCGCGCGCTTGGGTACGAGATG 518
QY 416 CACCGCTTGATGGGGGTCAACAGCACCGC 444
DB 517 CTCCCTTGTAGCTGACCGGAGCACGGC 489
```

## RESULT 15

```
US-10-023-523-48/c
: Sequence 48, Application US/10023523
: Patent No. US20020152485A1
: GENERAL INFORMATION:
: APPLICANT: Lees, Robert S.
: APPLICANT: Lees, Robert S.
: APPLICANT: Law, Simon W.
: APPLICANT: Arjona, Anibal A.
: TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
: TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
: FILE REFERENCE: 10797-004001
: CURRENT APPLICATION NUMBER: US/10/023, 523
: CURRENT FILING DATE: 2001-12-17
: PRIOR APPLICATION NUMBER: US/09/616, 289
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: US 09/517, 849
: PRIOR FILING DATE: 2000-03-02
: PRIOR APPLICATION NUMBER: US 08/979, 608
: PRIOR FILING DATE: 1997-11-26
: PRIOR APPLICATION NUMBER: US 60/031, 930
: PRIOR FILING DATE: 1996-11-27
: PRIOR APPLICATION NUMBER: US 60/048, 547
: PRIOR FILING DATE: 1997-06-03
: NUMBER OF SEQ ID NOS: 53
: SOFTWARE: FastSEQ for Windows Version 4.0
: SEQ ID NO 48
: LENGTH: 2561
: TYPE: DNA
: ORGANISM: Oryctolagus cuniculus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (246)...(1895)
US-10-023-523-48
```

Query Match 5.7%; Score 75.4; DB 12; Length 2561;

Best Local Similarity 53.0%; Pred. NO. 6e-06; Mismatches 181; Indels 2; Gaps 2;

```
QY 56 GCGCGCGCGCGCGCGCGCGCGCGCGCTGCGGCGGAAAGCTAGCGGAGAGCTCA 115
DB 875 GCGCGCGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 816
QY 116 GCGCGCGCGCGCGCGCGCGCGCGCTGCCAGCCATTTTCCGAGCCACCGCGGGCAC 175
DB 815 CGGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGG 756
QY 176 TCGCGAGCGCGCGCGGGGGTGCAGAGGAGAGCGCGGGGGGGCGAGCGAGCGGTCCG 235
DB 755 GCGCGGCTCGGGGGCGGGGCGAGGGGGCGCTGTGGCG-CCAGCGCGCGCGGGCGG 697
```

```
QY 236 CCGACTGAGCCCCCGGGCGCCCGGAACTTGGCGGGCGGAGCCCGCGAGCGGGG 295
DB 696 CCGGGGGCGCCCGGGGCGGGGCGGGGGCGGGCGAGCGGGCGGGCGGGGG 637
QY 296 CCGGCTTCCCGCGGGCGGCTCTGCATGCGGGGCCCACTCCGGGGCGGGCGGA 355
DB 636 CCGGGGTGGGCGGGCGGGGGCGGGCGGCG-GGCGGGGGCCCCCGGGGGGGCGGGCGGGG 578
QY 356 GCGCCCCCGGGCGCGCGAGCCCGCGCGCGCGCGCGCGCGCGCGCTCATG 415
DB 577 GCGGGCGGGGGTGGCTCCGGCGGGGGCGGCTGGACGGCGCGCGTTCGGTACGAGATG 518
QY 416 CACCGCTTGATGGGGGTCAACAGCACCGC 444
DB 517 CTCCCTTGTAGCTGACCGGAGCACGGC 489
```

Search completed: March 17, 2003, 18:33:21  
Job time : 126.606 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 11:34:54 ; Search time 2480.96 Seconds  
(without alignments)  
8623.366 Million cell updates/sec

Title: US-09-934-249-1  
Perfect score: 1321  
Sequence: 1 cgaccgcgcgtctgcgagcga.....ctgcgtagtgtaaaagcag 1321

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rnd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	859.4	65.1	1007	9	AL578575 AL578575
C 2	840.8	63.6	949	9	AL517150 AL517150
C 3	766.8	58.0	967	14	B0641849 B0641849
C 4	766	58.0	1046	14	BM922276 BM922276
C 5	736.4	55.7	916	14	BQ954555 BQ954555
C 6	726.6	55.0	805	9	AL558881 AL558881

Result	1	707	53.5	825	9	AL543170
LOCUS	AL578575/c	706.6	53.5	898	9	AL558882
DEFINITION	AL578575 LTL_NFL006.PL2 Homo sapiens cDNA clone CS0DK001YC24 3	630.6	47.7	1207	11	AK008976
ACCESSION	AL578575	618.6	46.8	1079	11	BC023092
VERSION	AL578575.1	614.6	46.5	782	14	BQ015170
KEYWORDS	EST	607.4	46.0	609	14	BQ636742
SOURCE	human.	578.4	43.8	890	14	BQ690750
ORGANISM	Homo sapiens	567.4	43.0	729	14	BQ575741
REFERENCE	1 (bases 1 to 1007)	564.8	42.8	730	14	BM677602
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	550	41.6	551	13	BM141979
TITLE	Full-length cDNA libraries and normalization	529.8	40.1	894	13	BI851941
JOURNAL	Unpublished (2001)	501.6	38.0	588	13	BM483503
COMMENT	Contact: Genoscope	493.6	37.4	693	9	AI761441
	Genoscope - Centre National de Sequencage	493.4	37.4	655	14	BO691705
	BP 191 91006 EVRY cedex - France	486	36.8	1400	13	BM559329
	Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.	468.8	35.5	1280	14	BO681500
	Location/Qualifiers	468	35.4	857	12	BG333347
	1. 1007	468	35.4	974	10	BB624904
	/organism="Homo sapiens"	468	35.4	763	13	BI646175
	/db_xref="taxon:9606"	468	35.4	626	14	BM974296
	/clone="CS0DK001YC24"	468	35.4	626	14	BO691066
	/clone_lib="LTL_NFL006.PL2"	468	35.4	626	14	BM712680
	/tissue_type="placenta"	468	35.4	626	14	BM712680
	/note="Vector: PCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end	468	35.4	626	14	BM712680

## ALIGNMENTS

RESULT 1  
LOCUS AL578575/c 1007 bp mRNA linear EST 16-FEB-2001  
DEFINITION AL578575 LTL\_NFL006.PL2 Homo sapiens cDNA clone CS0DK001YC24 3  
prime, mRNA sequence.  
ACCESSION AL578575  
VERSION AL578575.1 GI:12942781  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1007)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers  
1. 1007  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DK001YC24"  
/clone\_lib="LTL\_NFL006.PL2"  
/tissue\_type="placenta"  
/note="Vector: PCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com>

BASE COUNT 160 a 293 c 344 g 200 t 10 others  
ORIGIN

Query Match 65.1%; Score 859.4; DB 9; Length 1007;  
Best Local Similarity 97.7%; Pred. No. 4.2e-144;  
Matches 882; Conservative 9; Mismatches 10; Indels 2; Gaps 2;

```

OY 420 GCTTANTGGGGGTCACACGACCGCGCGCGCGCGCGCGCGAG-CCCAATCTCTCC 478
DB 1007 GCTTATGGGGGTCACACGACCGCGCGCGCGCGCGCGCGCAATTTCTCTCC 948
479 AGGTCAACTGCAACGCTCTTTGTTCCAGACATGAGATCAGGAGCTGGAGTTTGT 538
947 AGGTCAACTGCAACGCTCTTTGTTCCAGACATGAGATCAGGAGCTGGAGTTTGT 888
OY 539 CAGATCATCATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 598
DB 887 CAGATCATCATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 828
539 AGCCACTACAACTCTCTGACGCTCTTCATCAGCCGCGACAGCAGGAGGCGGAGAGA 658
DB 827 AGCCACMACAACTCTCTGACGCTCTTCATCAGCCGCGACAGCAGGAGGCGGAGAGA 768
OY 659 CAGATGCTCTCTCTCTCAGAGAGATGCTGTGGCTGTGAGAGACAGTGTACAGCAAC 718
DB 767 GAGATGCTCTCTCTCTCAGAGAGATGCTGTGGCTGTGAGAGACAGTGTACAGCAAC 708
719 GGAATCCAGAGCGGAGGCTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 778
OY 707 GGAATCCAGAGCGGAGGCTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
779 CCCTTGGCCAGCGGAGGCTTCCAGCGCTTCCAGCCACTATCCGTACTGACAGC 838
DB 648 CCCTTGGCCAGCGGAGGCTTCCAGCGCTTCCAGCCACTATCCGTACTGACAGC 589
OY 839 GAGATGACCTGCTCCACCACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 898
DB 588 GAGATGCTCTCTCTCTCAGAGAGATGCTGTGGCTGTGAGAGAGAGAGAGAGAG 529
899 GAGATGCTCTCTCTCTCAGAGAGATGCTGTGGCTGTGAGAGAGAGAGAGAGAG 958
OY 528 AGCCCTGACCGCTCTCAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 469
959 GTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1018
DB 468 GTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 409
OY 1019 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1078
DB 408 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 349
1079 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1138
DB 348 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 289
1139 TCTTTCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1198
DB 288 TCTTTCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 229
OY 1199 CACACACACATCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1258
DB 228 MACACACACATCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 169
OY 1259 AAAGAGACCGCTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1318

```

DB 168 AAAGAACCCCTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 109  
OY 1319 CAG 1321  
DB 108 CAG 106

RESULT 2  
LOCUS AL517150/c 949 bp mRNA linear EST 13-FEB-2001  
DEFINITION AL517150 LTI\_NFL011.NBC1 Homo sapiens cDNA clone CS0DA008YB23 3  
prime, mRNA sequence.  
ACCESSION AL517150  
VERSION AL517150.1 GI:12780643  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 949)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

FEATURES  
Source  
1..949  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DA008YB23"  
/clone\_1lb="LTI\_NFL011.NBC1"  
/sex="male"  
/tissue\_type="neuroblastoma cells"  
/lab\_host="DH10B"  
/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library is not normalized, but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com>"

BASE COUNT 145 a 289 c 330 g 183 t 2 others  
ORIGIN

Query Match 63.6%; Score 840.8; DB 9; Length 949;  
Best Local Similarity 98.8%; Pred. No. 8.8e-141;  
Matches 856; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

```

OY 456 CCGGGCAGCCCAATGCTCTCTGACGCTCACTGCAACGCTCTTTGTTCCAGAGCATGG 515
DB 949 CCGGGCAGCCCAATGCTCTCTGACGCTCACTGCAACGCTCTTTGTTCCAGAGCATGG 890
OY 516 AGATACGAGAGTGAGTTGTTGATCATCATCATCATCATCATCATCATCATCATCATCAT 575
DB 889 AGATACGAGAGTGAGTTGTTGATCATCATCATCATCATCATCATCATCATCATCATCAT 830
OY 576 TGTGTGATCATCATGCTCTGAGCCACTCAAGAGCTCTGACAGGCTCTTCATCAGCC 635
DB 829 TGTGTGATCATCATGCTCTGAGCCACTCAAGAGCTCTGACAGGCTCTTCATCAGCC 770
OY 636 GGCACAGCCAGGGGGGAGAGAGAGATGCTCTCTCAAGAGATGCTGTGGCCT 695
DB 769 GGCACAGCCAGGGGGGAGAGAGAGATGCTCTCTCAAGAGATGCTGTGGCCT 710
OY 696 CGGAGAGCAGAGTGTACAGCAACGGAATCCAGAGCCGACAGTGTACGCCCGGCTGGC 755
DB 709 CGGAGAGCAGAGTGTACAGCAACGGAATCCAGAGCCGCA-GTCTACGCCCGGCTGGC 651

```

QY	756	CCACCGACCGCGTGGCGCGTGGCGCCCTTTCGGCCACGCGGAGGCGCTTCCACCGCTTCCAGC	815
Db	650	CCACCGACCGCGTGGCGCGTGGCGCCCTTTCGGCCACGCGGAGGCGCTTCCACCGCTTCCAGC	591
QY	816	CCACCTATCCGCTACTCGCAGCAGCAGATCGACTTGCACCCACACATCTCGCTGCAGACG	875
Db	590	CCACCTATCCGCTACTCGCAGCAGCAGATCGACTTGCACCCACACATCTCGCTGCAGACG	531
QY	876	GGGAGGACCCCCACCCCTTACCAAGGCCCCCTGCACCTTCCAGCTTGGGACCCCGACGC	935
Db	530	GGGAGGACCCCCACCCCTTACCAAGGCCCCCTGCACCTTCCAGCTTGGGACCCCGACGC	471
QY	936	AGCTGAACCTGAACCGGGAGTGGTGGCGGACCCCAACAGAACATCTTCGACAGTG	995
Db	470	AGCTGAACCTGAACCGGGAGTGGTGGCGGACCCCAACAGAACATCTTCGACAGTG	411
QY	996	ACCTGATGATATAGTCCACGAGCTGGGCGGCCCCCTGCCCCCGACAGTAACTCGGCGATCA	1055
Db	410	ACCTGATGATATAGTCCACGAGCTGGGCGGCCCCCTGCCCCCGACAGTAACTCGGCGATCA	351
QY	1056	GGCGCAGCTGCTACGCGCAGCGGCGGGCGCATGAGAGGGCGCGCCACCTTACAGCGAGG	1115
Db	350	GGCGCAGCTGCTACGCGCAGCGGCGGGCGCATGAGAGGGCGCGCCACCTTACAGCGAGG	291
QY	1116	TCATGCGGCACCTACCGCGGGGTCTCTCTTCGACACACAGACAGACGATGGCGCCCTCT	1175
Db	290	TCATGCGGCACCTACCGCGGGGTCTCTCTTCGACACACAGACAGACGATGGCGCCCTCT	231
QY	1176	TGCTTGAGGGGAGCCCGGCTCCACACACACATCGGCCCTTAGAGAGCGAGCATCT	1235
Db	230	TGCTTGAGGGGAGCCCGGCTCCACACACACATCGGCCCTTAGAGAGCGAGCATCT	171
QY	1236	GGAGCAAAAGAGAGATTAACAGAAAGGACACCTCTTAGGGTCCCAAGGGGGCGCGG	1295
Db	170	GGAGCAAAAGAGAGATTAACAGAAAGGACACCTCTTAGGGTCCCAAGGGGGCGCGG	111
QY	1296	CTGGGGCTGCGGTAGGTGAAAAAGGCGAG	1321
Db	110	CTGGGGCTGCGGTAGGTGAAAAAGGCGAG	85
RESULT 3			
LOCUS	B0641849	967 bp	mRNA linear EST 15-JUN-2002
DEFINITION	AGENCOURT_8287174 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6292265		
ACCESSION	B0641849		
VERSION	B0641849.1	GI:21766021	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 967)		
JOURNAL	NIH-MGC <a href="http://mhc.nci.nih.gov/">http://mhc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a>		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Rubin Laboratory		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>		
	Plate: L1CM2493 row: 9 column: 18		
	High quality sequence stop: 571.		
FEATURES	Location/Qualifiers		
Source	1. 967		
	"/organism="Homo sapiens"		
	"/db_xref="taxon:9606"		

[illegible]

[illegible]

pcmv="Organ: pooled brain, lung, testis; Vector: PCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library."

Query Match	58.0%;	Score 766;	DB 14;	Length 1046;
Best Local Similarity	95.8%;	Pred. No. 2.1e-127;		
Matches 818; Conservative	0;	Mismatches 32;	Indels 4;	Gaps 3

Qy	347	CGGGCCGAGACCCCCCGGGGGGCCGCCGAGCCCCCGGGGGCCCGGCSCGCGCGCGCG	406
Dd	1	CCGGCCGGAGCCCCCCCCGGGGCCCCCGGAGCCCCCGGGCCCCCGCCGCCGCGC	60
Qy	407	CGGTCAATCACCGCTTGATGGGGGTCAACAGACCGCGCGCGCGCGCGCGGACAACC	466
Dd	61	CGTCCATTCACACCGCTTGATGGGGGTCAACAGACCGCGCGCGCGCGCGCGGACAACC	120
Qy	467	AATGTCCTCGACAGTCGAACCTCTTTGTTCCAGACCATGGAGATCAGGAG	526
Dd	121	AATGTCCTCGACAGTCGAACCTCTTTGTTCCAGACCATGGAGATCAGGAG	180
Qy	527	CTGGAGTTGTTCAATCATCATCTGCGTGGTGGATGAGTGGATGGTGGTATC	586
Dd	181	CTGGAGTTGTTCAATCATCATCTGCGTGGTGGATGAGTGGATGGTGGTATC	240
Qy	587	ACGTGCTCTGAGCCACACAAAGCTGCTGACCGGTCTTCATCAGCCGGACAGCCAG	646
Dd	241	ACGTGCTCTGAGCCACACAAAGCTGCTGACCGGTCTTCATCAGCCGGACAGCCAG	300
Qy	647	GGGCGAGAGAGAAAGATGCCCTGTCTCAGAAGGATGCTGTGGCCCTCGGAGACACA	706
Dd.	301	GGGCGAGAGAGAAAGATGCCCTGTCTCAGAAGGATGCTGTGGCCCTCGGAGACACA	360

OY	707	GTGTACAGCAACGGAATATCCAGAGACCGCAGATCTACAGCCCCCGCTGGCCACACGACCG	766
Db	361	GTGTACAGCAACGGAATATCCAGAGACCGCAGATCTACAGCCCCCGCTGGCCACACGACCG	420
OY	767	CTGAGCCGTCGCCCTTCGCGCCAGCGGAGCGCTTTCACAGCCCTTCACGCCACTATTCG	826
Db	421	CTGGGCCGTGGCCCTTCGCGCCAGCGGAGCGGCTTCACAGCCCTTCACGCCACTATTCG	480
OY	827	TACCTGCAGACACAGATTCGACCTGCACCCACACATTCGCTGTAGACGGGAGAGACCC	886
Db	481	TACCTGCAGACACAGATTCGACCTGCACCCACACATTCGCTGTAGACGGGAGAGACCC	540
OY	887	CCACCTCTACCAAGGGGCCCTTCGACCCCTCCAGCTTCGGGAGCCCGAGACAGCTGGAACTG	946
Db	541	CCACCTCTACCAAGGGGCCCTTCGACCCCTCCAGCTTCGGGAGCCCGAGACAGCTGGAACTG	600
OY	947	AACCGGAGATGGTGTGCGCGCACCCCAACAAAGAACATCTTCGACAGTAGACCTATGGAT	1006
Db	601	AACCGGAGATGGTGTGCGCGCACCCCAACAAAGAACATCTTCGACAGTAGACCTATGGAT	660
OY	1007	AGTGCAGAGCTGTGGCGGGCCCTTCGCCCCCGACAGCACTACTGGGCACTATGCGGCCACTGC	1066
Db	661	AGTGCAGAGCTGTGGCGGGCCCTTCGCCCCCGACAGCACTACTGGGCACTATGCGGCCACTGC	720
OY	1067	TACGGCAGACGGCGG--GCGCATATGAGAGGGGCGCGCG--CCACCTACAGCAGATTCATCGGCC	1124
Db	721	TACGGCAGAGGGGCGCGCGCATATGAGAGGGGCGCGCGCCCACTACAGGAGATTCATCGGCC	780
OY	1125	ACTTA--CCCGGGGATCTCTCTTCACAGCAGCAGCAGAGCAGAGTGGGCGCCCTCTCTTCTGGA	1182
Db	781	ACTTAACCCGGGGGATCTCTCTTCACAGCAGCAGCAGCAGTGGGCGCGCCCTCTCTTCTGCG	840
OY	1183	GGGAGACCGCGCTTC	1196
Db	841	CTGGAGAGGGGAAC	854

RESULT	5
LOCUS	B0954555
DEFINITION	BQ954555 916 bp mRNA linear EST 21-AUG-2002
ACCESSION	BQ954555
VERSION	BQ954555
KEYWORDS	IMAGE:6204609 5', mRNA sequence.
SOURCE	BQ954555.1 GI:22370033
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue Procurement: Dr. James R. Lupski CDNA library preparation: Life Technologies, Inc. CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM13626 row: c column: 10 High quality sequence stop: 669. Location/Qualifiers 1..916
FEATURES	
source	

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6204609"
/lupski_lib="lupski_sciatic_nerve"
/sex="male"
/tissue_type="sciatic nerve"

```



|||||  
Db 445 CCGGAGTGGTGGCGGACCCCAACAGACCATTTTGACAGTACGTATGATAG 386  
QY 1009 TGGCAGGCTGGGCGGCGCTCCCGCCAGCATGTAAGGCGCCACAGTCTA 1068  
Db 385 TGGCAGGCTGGGCGGCGCTCCCGCCAGCATGTAAGGCGCCACAGTCTA 326  
QY 1069 CGGACGCGGCGGCGCATGAGAGGGCGCGCCAGCATGAGAGGTCATCGGCACTA 1128  
Db 325 CGGACGCGGCGGCGCATGAGAGGGCGCGCCAGCATGAGAGGTCATCGGCACTA 266  
QY 1129 CGGCGGCTGCTCTCTCCAGACACAGAGCATGAGAGGCGCGCTCTCTGAGAGGAC 1188  
Db 265 CGGCGGCTGCTCTCTCTCCAGACACAGAGCATGAGAGGCGCGCTCTCTGAGAGGAC 206  
QY 1189 CGGCGTCCACACACACATGCGCGCTCTGAGAGGCGCGCATCTGAGAGAGAA 1248  
Db 205 CGGCGTCCACACACACATGCGCGCTCTGAGAGGCGCGCATCTGAGAGAGAA 146  
Db 1249 GGTAAACAGAAAGACACCTCTAGAGGTCGCCAGGCGGCGCGCTGAGAGT 1308  
Db 145 GGTAAACAGAAAGACACCTCTCTAGAGGTCGCCAGGCGGCGCGCTGAGAGT 86  
QY 1309 GGTAAAGGACAG 1321  
Db 85 GGTAAAGGACAG 73

RESULT 7  
AL543170 825 bp mRNA linear EST 16-FEB-2001  
LOCUS AL543170 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1002Y103 5  
DEFINITION AL543170 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1002Y103 5  
ACCESSION AL543170  
VERSION AL543170.1 GI:12875648  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 825)  
AUTHORS L.W.B., Gruber,C., Jesse,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
ATURES Location/Qualifiers  
source 1..825  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1002Y103"  
/clone\_1bp="LTI\_NFL006.PL2"  
/tissue\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email: fliang@lifestech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 125 a 327 c 229 g 128 t 16 others  
ORIGIN

Query Match 53.5%; Score 707; DB 9; Length 825;  
Best Local Similarity 92.6%; Pred. No. 7.3e-117;  
Matches 760; Conservative 13; Mismatches 44; Indels 4; Gaps 3;  
QY 110 GGCTCAGCCCGGCGGCGCGCGCGCTGCGCAGCCATTTCGAGCCGACCCGC 169

|||||  
Db 1 GGCTCAGCCCGGCGGCGCGCGCGCGCTGCGCAGCCATTTCGAGCCGACCCGC 60  
QY 170 GGGCACTGCGCGAGCCCGCGGCGCTGCGAGAGGCGCGGCGGCGCGAGCGGCG 229  
Db 61 GGGCACTGCGCGAGCCCGCGGCGCTGCGAGAGGCGCGGCGGCGCGAGCGGCG 120  
QY 230 GTCCCGCGACTGAGCCCGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCG 289  
Db 121 GTCCCGCGACTGAGCCCGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCG 180  
QY 290 CGGAGGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 348  
Db 181 CGGAGGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
QY 349 GCGCGAGCCCG 408  
Db 241 GCGCGAGCG 300  
QY 409 GTCCATGACACCGCTGATGGGGTCAGACCGCGCGCGCGCGCGCGCGCGCGCG 468  
Db 301 TTCATGACACCGCTTATTTGGGTCAAGACCGCGCGCGCGCGCGCGCGCGCGCA 360  
QY 469 TGTCTCTGAC--GTGCACTGCAAGCGCTTTTGTTCAGAGCATGAGATGACGAG 526  
Db 361 TTCTCTCTGACCTGCTCTCAATTCAGACCTTTTTCATGACATGAGATGACGAG 420  
QY 527 CTGAGATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 586  
Db 421 CTGAGATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
QY 587 ACGTCTGTGAGCCACTGACAGCTGTGACAGGCTCTTATGACCGCGCGAGCGCG 646  
Db 481 ACGTCTGTGAGCCACTGACAGCTGTGACAGGCTCTTATGACCGCGCGAGCGCG 540  
QY 647 GGGGAGAGAGAGATGCGCTGCTCTGAGAGATGCTGCGCGCTGCGAGAGAC 706  
Db 541 GGGGAGAGAGAGATGCGCTGCTCTGAGAGATGCTGCGCGCTGCGAGAGAC 600  
QY 707 GTGTCAGGCAAGGATGCGAGAGCGCGAGAGCTTACGCGCGCGCTGCGAGAGCG 766  
Db 601 GTGTCAGGCAAGGATGCGAGAGCGCGAGAGCTTACGCGCGCGCTGCGAGAGCG 660  
QY 767 CTGCGCGTGGCGCGCTTGGCCAGCGGAGGCGCTTCCAGCGCTTCCAGCGCTTCC 826  
Db 661 CTGCGCGTGGCGCGCTTGGCCAGCGGAGGCGCTTCCAGCGCTTCCAGCGCTTCC 720  
QY 827 TACCTGACAGAGATGACCTGCGACCGACCATCTGCTGTGACAGCGGAGAGCC 886  
Db 721 TACCTGACAGAGATGACCTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 779  
QY 887 CCACCTACAGAGGCGCGCTGACCGCTTCCAGCTTGGGAGCC 927  
Db 780 CCACCTACAGAGGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 820

RESULT 8  
AL558882 898 bp mRNA linear EST 16-FEB-2001  
LOCUS AL558882 LTI\_NFL008.TC2 Homo sapiens cDNA clone CS0J015YF12 5  
DEFINITION AL558882 LTI\_NFL008.TC2 Homo sapiens cDNA clone CS0J015YF12 5  
ACCESSION AL558882  
VERSION AL558882  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 898)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope



Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [sequef@genoscope.cns.fr](mailto:sequef@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

## FEATURES

Location/Qualifiers  
source

1. 898  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DJ015F12"  
/clone\_11b="LTI\_NFL008\_Tc2"  
/sex="male"  
/issue\_type="T cells from T cell leukemia"  
/note="Vector: PCWSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCWSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL : <http://fulllength.invitrogen.com>"

BASE COUNT 196 a 310 c 260 g 126 t 6 others  
ORIGIN

Query Match 53.5%; Score 706.6; DB 9; Length 898;  
Best Local Similarity 99.3%; Pred. No. 8.6e-117;  
Matches 726; Conservative 2; Mismatches 1; Indels 2; Gaps 2;

QY 589 GTGCTGCTGAGCCATACAGCTGTCTGACGGTCTTTCATCAGCCGCGACAGCCAGG 648  
D 1 GTGCTGCTGAGCCATACAGCTGTCTGACGGTCTTTCATCAGCCGCGACAGCCAGG 60  
QY 649 GCGAGAGAGAAAGTCCCTGCTCTCAGAAAGATGCTGTGGCCCTTGAGAGCAGT 708  
D 61 GCGAGAGAGAAAGTCCCTGCTCTCAGAAAGATGCTGTGGCCCTTGAGAGCAGT 120  
QY 709 GTGAGGCAAGGAAATCCAGAGCCGAGAGCTTACGCCCCGCTGCGCCAGCAGCCGCT 768  
D 121 GTGAGGCAAGGAAATCCAGAGCCGAGAGCTTACGCCCCGCTGCGCCAGCAGCCGCT 179  
QY 769 GGGCGTGGCCGCTTGGCCGAGCGGGAGCGCTTCCAGCCGCTTCCAGCCCATATCCGTA 828  
D 180 GGGCGTGGCCGCTTGGCCGAGCGGGAGCGCTTCCAGCCGCTTCCAGCCCATATCCGTA 239  
QY 829 CCTCAGACAGCAGATGCACTGCGCCAGCCAGCATCTCGTGTGACAGCGGGAGAGCCGCC 888  
D 240 CCTCAGACAGCAGATGCACTGCGCCAGCCAGCATCTCGTGTGACAGCGGGAGAGCCGCC 299  
QY 889 ACCCTACAGAGGCGCCCTGACCTTCAGACTTCGGGAGCCCGAGCAGCAGTGGAACTGAA 948  
D 300 ACCCTACAGAGGCGCCCTGACCTTCAGACTTCGGGAGCCCGAGCAGCAGTGGAACTGAA 359  
QY 949 CCGGAGTGTGGTGGCGGACCCCAAGAGAACCATCTTTCAGACAGTGAATGATAG 1008  
D 360 CCGGAGTGTGGTGGCGGACCCCAAGAGAACCATCTTTCAGACAGTGAATGATAG 419  
QY 1009 TGCCAGGTGG 1068  
D 420 TGCCAGGTGG 479  
QY 1069 CGGAGAGCGGG 1128  
D 480 CGGAGAGCGGG 539  
QY 1129 CCGGGGGGTCTCTCTTCAGACAGCAGAGAGTGGGGGGGGGGGGGGGGGGGGGGGGGG 1188  
D 540 CCGGGGGGTCTCTCTTCAGACAGCAGAGAGTGGGGGGGGGGGGGGGGGGGGGGGGGG 599  
QY 1189 CCGGGTCCACACACACATCGCGCCCTTTCAGAGAGCCAGCCATCTGGAGCAAGAGAA 1248  
D 600 CCGGGTCCACACACACATCGCGCCCTTTCAGAGAGCCCA-SCATCTTCAGAGCAAGAGAA 658  
QY 1249 GGATTAACAGAAAGACACCTCTCTAGGGTCCCAAGGGGGGGGGGGGGGGGGGGGGGG 1308

Db 659 GGATTAACAGAAAGACACCTCTCTAGGGTCCCAAGGGGGGGGGGGGGGGGGGGGGGGTA 718  
QY 1309 GGTGAAAAAGCGAG 1321  
Db 719 GGTGAAAAAGCGAG 731

## RESULT 9

AK008976  
LOCUS

## DEFINITION

AK008976 1207 bp mRNA linear HTC 19-JAN-2002  
Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210418102.Nedd4 WW binding protein 4, full insert sequence.

## ACCESSION

AK008976  
VERSION  
KEYWORDS  
SOURCE

## ORGANISM

Mus musculus

## REFERENCE

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

## AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
10421259

## TITLE

JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishikawa, K., Kitahara, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishimura, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matsumoto, Y., Onoda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawaji, Y., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.  
RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichipillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## TITLE

JOURNAL  
MEDLINE  
PUBMED

## AUTHORS

Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nkaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamly, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mommaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Sessa, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660

PUBMED  
11217851

5 (bases 1 to 1207)

Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,  
Arikawa,T., Balderelli,R., Bono,H., Brownstein,M., Butl,C.,  
Carinini,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanaoka,T.,  
Hara,A., Hayatsu,N., Hill,I.Y., Hiramoto,K., Hiroka,T., Horii,F.,  
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,  
Kato,H., Kawai,J., Koijima,Y., Konno,H., Kouda,M., Koyama,  
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,  
Numazaki,K., Ohno,M., Okazaki,Y., Okido,T., Owe,C., Quackenbush,J.,  
Salto,H., Salto,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,  
Schmidt,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,  
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,  
Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and  
Hayashizaki,Y.

Direct Submission

Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-resgsc.riken.go.jp,  
URL:http://genome.resc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGGACGAGATCCAGAAGCTCTTTTATTCTTTTAA 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of normalization to Rot = 5.0 and subtraction Rot  
= 25.0. Second strand cDNA was prepared with the primer adapter of  
sequence 15'  
GAGGACGAGATTCGCAGTAATTAATTAATGAATCCCCCCCCC 3'. cDNA was cleaved  
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.  
Host: lambda.

Location/Qualifiers

1..1207

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="RANTOM,DB:2210418102"  
/db\_xref="MGD:MGI:1992457"  
/db\_xref="taxon:10090"  
/clone="2210418102"  
/sex="male"  
/tissue-type="stomach"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
391..1173  
/note-"Need4 WW binding protein 4  
data source:MGD, source key:MGI:1929600, evidence:ISS  
putative"  
/codon\_start=1  
/protein\_id="BAB26001.1"  
/db\_xref="GI:12843488"  
/db\_xref="MGD:MGI:1929600"  
/translation="MSFKRAKAQSRSLFSPSEITELFEVQIVLVVMVMVMTCLL  
SHVLASAFYSIRHSQAARRDDGLSSGGCWLPSSVSYGKPEQVYAAPPDRDLAVV  
PPFLQRSFPQTPTLQHEIALPPTILSISSGEPPPYQGCTTLRDPEOLELRNS  
VRAPNTIDISLDISTMTGSGPCPPSSNIGISATCYSSGGRMGPPRPPTYSEVIGHNP  
GSSPHOOSNPSLSLEGTGRHNHSHPLENKEKKRKGPL"

BASE COUNT 200 a 459 c 374 g 173 t 1 others

ORIGIN

Query Match 47.7%; Score 630.6; DB 11; Length 1207;  
Best Local Similarity 82.4%; Pred. No. 3.3e-103;  
Matches 782; Conservative 0; Mismatches 139; Indels 28; Gaps 4;

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (04-FEB-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarathne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 36 Row: h Column: 6  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Similarity but not  
identity to protein  
This clone has the following problem: frame shifted.  
Location/Qualifiers  
1..1079  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="CZEBCH II"  
/clone="IMAGE:3989996"  
/tissue\_type="Mammary tumor metastatized to lung. Tumor  
arose spontaneously from a senescent normal mammary  
(clonal) outgrowth infected with the virus MMTV."  
/clone\_1id="NCI\_CGAP\_Lu29"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
BASE COUNT 271 a 341 c 289 g 178 t  
ORIGIN  
Query Match 46.8%; Score 618.6; DB 11; Length 1079;  
Best Local Similarity 84.5%; Pred. No. 4.6e-101;  
Matches 741; Conservative 0; Mismatches 109; Indels 27; Gaps 3;  
QY 417 ACCGCTTGATGGGGTCAACAGCACCGCGCGCGCGGGGAGCCCAATGTCTCT 476  
|||||  
Db 15 ACCCTTGATGGGGTCAACAGCACCGCGCGCGCGGGGAGCCCAATGTCTCTCT 74  
|||||  
QY 477 GCAGGTGAATCGTGAACGCTCTTGTTCAGAGCATGAGATCAGGAGTGTG 536  
|||||  
Db 75 GCGGGTGAATCGTGAACGCTCTTGTTCAGAGCATGAGATCAGGAGTGTG 134  
|||||  
QY 537 TTCAGATCATCATGT 596  
|||||  
Db 135 TGCAGATCATCATGT 194  
|||||  
QY 597 TGAGCCACTACAGCTGTCTCAGCGTCTTCATCAGCCGACAGCCAGGCGGAGGA 656  
|||||  
Db 195 TGAGCCACTACAGCTGTCTCAGCGTCTTCATCAGCCGACAGCCAGGCGGAGGA 254  
|||||  
QY 657 GAGAAGATGCCCTGCTCAGAGAGATGCTGTGGCCCTGGAGAGCACTATGTCAGCA 716  
|||||  
Db 255 GAGAGATGAGCTGTCTCAGAGAGATGCTGTGGCCCTGGAGAGTACGCTGTACAGG-- 312  
|||||  
QY 717 ACGGATTCACAGAGCCGAGGTCTACGCGCGCTCGGCGCCACGACCGCTGCGCTGTGC 776  
|||||

Db 313 -TGAATGCCGAGACACAGAGTCTATGCCCCCGCTCGGCCCACTGACCGACTGCTGTGC 371  
QY 777 CGCCCTTCGCCAGCGGAGGCGTTCACACCGCTTCACGCCCACTATCCGACTGTGAGC 836  
|||||  
Db 372 CCCCCTTCATCCAGCG-----ACCGATTCCAAACCACTATCCCTGACGAC 422  
|||||  
QY 837 ACAGATTCAGACTCCACACCCACATCTCGCTGTGAGAGGAGAGAGGAGGAGGAGGAGG 896  
|||||  
Db 423 ACGAATTCGCTGTGACACCCACATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 482  
|||||  
QY 897 AGGCGCCCTGACCCCTCCAGCTTGGGAGCCCGAGCAGCAGCTGTGAATGAACTGAGCGGAGT 956  
|||||  
Db 483 AGGCGCCCTGACCCCTCCAGCTTGGGAGCCCGAGCAGCAGCTGTGAATGAACTGAGCGGAGT 542  
|||||  
QY 957 CGGTGGCGGAGCCCGCAACAGAACATCTTCGACAGGAGGAGGAGGAGGAGGAGGAGGAGG 1016  
|||||  
Db 543 CTGTGGCGGAGCCCGCAACAGAACATCTTCGACAGGAGGAGGAGGAGGAGGAGGAGGAGG 602  
|||||  
QY 1017 TGGCGGCGCCCTGCGCCCGCCAGAGTAATCGGAGCATGAGGAGGAGGAGGAGGAGGAGGAGG 1076  
|||||  
Db 603 TGGGAGGCGCCCTGTCTCCCGCCAGAGTAATCGGAGCATGAGGAGGAGGAGGAGGAGGAGG 662  
|||||  
QY 1077 GCGGAGCGCATGAGAGGAGGCGCCCGCCACCTACAGCAGAGTATGCGCATACCGGAGGT 1136  
|||||  
Db 663 GTGGGCGCATGAGAGGAGGCGCCCGCCACCTACAGCAGAGTATGCGCATACCGGAGGT 722  
|||||  
QY 1137 CCTCTTCACACACAGAGAGAGTGTGGGCGCCCTCTCTCTGTGAGGAGGAGGAGGAGGAGG 1196  
|||||  
Db 723 CCTCTTCACACACAGAGAGAGTGTGGGCGCCCTCTCTCTGTGAGGAGGAGGAGGAGGAGG 782  
|||||  
QY 1197 ACCACACACATGCGCGCCCTTGTGAGAGCGAGCATGTGAGAGAGAGAGATTAAC 1256  
|||||  
Db 783 ATCACTGCGACATGTGCCCACTGGA-----GAACAAGAGAGAGAGAGAAC 827  
|||||  
QY 1257 AGAAGAGACACCTCTGTAGGCTCCCGAGGAGGAGCGG 1293  
|||||  
Db 828 AGAAGAGTCAACCCCTCTAGAGAGTGGGAGCGGAGGCG 864  
|||||  
RESULT 11  
B0015170/c 782 bp mRNA linear EST 26-MAR-2002  
LOCUS  
DEFINITION  
UI-H-ED1-axw-k-20-0-UI-s1 NCI\_CGAP\_ED1 Homo sapiens CDNA clone  
IMAGE:5834635 3', mRNA sequence.  
ACCESSION  
B0015170  
VERSION  
B0015170.1 GI:19740071  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 782)  
AUTHORS  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Dr. Jose Mercuende  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Seq primer: M13 FORWARD  
POLYA=Yes.  
FEATURES  
Location/Qualifiers  
1..782  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5834635"  
/clone\_1id="NCI\_CGAP\_ED1"  
/tissue\_type="Chondrosarcoma"

```

/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/notes="Organ: Left Pubic Bone; Vector: pT73-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI-GAP-ED1 is a normalized cDNA library
containing the following tissue(s): Chondrosarcoma cell
line C5. The library was constructed according to Bonaldo
, Lennon and Soares, genome Research, 6:791-806, 1996.
First strand cDNA synthesis was primed with an oligo-dT
ligated to an EcoR I adaptor, digested with Not I, and
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (GT)18 tail. The
sequence tag for this library is GCTCAAGGCT.
TAG LIB-UI-H-ED1
TAG TISSUE=chondrosarcoma
TAG_SEQ=GCTCAAGGCT"

BASE COUNT      109 a      223 c      271 g      176 t      3 others
IGIN

Query Match      46.5%; Score 614.6; DB 14; Length 782;
Best Local Similarity 98.9%; Pred. No. 2.4e-100;
Matches 617; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 698 GAGAGCAGCTGTCAGGACAGGAAATCCAGAGCGAGTCTACGCCCGCTCGGCC 757
Db 782 GAGAGCAGCTGTCAGCAAAAGAAATCCAGAGCGAGTCTACGCCCGCTCGGCC 723
QY 758 ACCGACCGCTGCGCGCGCTGCGCGCGAGGAGGCGCTTCACCGCTTCACGCC 817
Db 722 ACCGACCGCTGCGCGCTGCGCGCGAGGAGGCGCTTCACCGCTTCACGCC 663
QY 818 ACCTATCGTACCTGACAGAGATGCACTGCGCACCAACATCTGCTGACAGAGG 877
Db 662 ACCTATCGTACCTGACAGAGATGCACTGCGCACCAACATCTGCTGACAGAGG 603
QY 878 GAGAGCGCGCGCGCTGACAGGCGCGCTGACAGGCGCGCGCGCGCGCGCGCG 937
Db 602 GAGAGCGCGCGCGCGCTGACAGGCGCGCTGACAGGCGCGCGCGCGCGCGCG 543
QY 938 CTGGAACGAGCGGAGTGGTGGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCG 997
Db 542 CTGGAACGAGCGGAGTGGTGGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCG 483
QY 998 CTGATGATAGTGGCAGGCTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1057
Db 482 CTGATGATAGTGGCAGGCTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 423
QY 1058 GCCACGTCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1117
Db 422 GCCACGTCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 363
QY 1118 ATGGGCGACTACCGCGGCGCTCTCTCCAGCAGCAGAGAGAGTGGCGCGCTG 1177
Db 362 ATGGGCGACTACCGCGGCGCTCTCTCCAGCAGCAGAGAGAGTGGCGCGCTG 303
QY 1178 CTGGAGGAGACCGCGCTCCACACACACATCGCGCGCTAGAGAGGCGAGCATGG 1237
Db 302 CTGGAGGAGACCGCGCTCCACACACACATCGCGCGCTAGAGAGGCGAGCATGG 243
QY 1238 AGCAAAAGAGAGATAAACAGAAAGACACCTCTCTAGGGTCCCGAGGGGCGCGGCT 1297
Db 242 AGCAAAAGAGAGATAAACAGAAAGACACCTCTCTAGGGTCCCGAGGGGCGCGGCT 183
QY 1298 GGGGCTGCTAGGTAAAGAGCG 1321
Db 182 GGGGCTGCTAGGTAAAGAGCG 159

```

RESULT 12  
B0636742

```

LOCUS      B0636742      609 bp      mRNA      linear      EST 15-JUL-2002
DEFINITION hd13h06.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
ACCESSION B0636742
VERSION   B0636742
KEYWORDS  B0636742.1 GI:21761201
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 609)
AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman
,J.W., Bouffard,G., Smith,D. and Peterson,K.
Expressed sequence tag analysis of human retina for the NEIRBank
project. Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
CONTACT:  Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: grawemehellix.nih.gov
Plate: 13 row: h column: 06
Seg primer: M13Rpl reverse primer (ABI).
Location/Qualifiers
1..609
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hd13h06"
/clone_lib="Human Retina cDNA (Un-normalized, unamplified)
): hd/he"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/notes="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-pGACTGATCTAGATCGGCGCGCGCGCG(T)15-3'
]. EST analysis was performed on the unamplified library
at the NIH Intramural Sequencing Center (NISC)."
```

```

BASE COUNT      114 a      238 c      182 g      75 t
ORIGIN

Query Match      46.0%; Score 607.4; DB 14; Length 609;
Best Local Similarity 99.8%; Pred. No. 4.7e-99;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 631 CAGCGGCAAGCAGGCGGCGGAGAGAAAGATCCCTGTCCTCAGAGATGCTGTG 690
Db 1 CAGCGGCAAGCAGGCGGCGGAGAGAAAGATGCTGTCCTCAGAGATGCTGTG 60
QY 691 GCCCTGGAGACAGATGTCAAGGCAAGTCCAGAGCGCGAGGTCTTACGCCCGCC 750
Db 61 GCCCTGGAGACAGATGTCAAGGCAAGTCCAGAGCGCGAGGTCTTACGCCCGCC 120
QY 751 TCGGGCCACCGACCGCTGCGCGCGCGCTTGGCCACGCGGAGGCGCTTCAACGCTT 810
Db 121 TCGGGCCACCGACCGCTGCGCGCGCGCTTGGCCACGCGGAGGCGCTTCAACGCTT 180
QY 811 CCAGGCCACCTATCGCTGACAGACGAGATGACCTGCCACCACTCTCGCTGTC 870
Db 181 CCAGGCCACCTATCGCTGACAGACGAGATGACCTGCCACCACTCTCGCTGTC 240
QY 871 AGAGGGGAGGAGCGCCGACCGCTTACGAGGCGCGCTGACCGCTTCAAGCTTGGGAGCCCGCA 930

```



DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, bento-soares@uiowa.edu  
 Seq primer: M13 FORWARD  
 POLYA-Yes

# FEATURES

## SOURCE

Location/Qualifiers

1. 729

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-H-E21-bdg-h-14-0-UI"

/tissue="Chondrosarcoma Grade II"

/dev\_stage="Adult"

/lab\_host="DH10B (Life Technologies)"

/note="Organ: Left Pelvis; Vector: pRT73-Pac (Pharmacia)

with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I;

NCI.GAP.Ch2 is a normalized cDNA library containing the

following tissue(s): Chondrosarcoma Grade II. The library

was constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pRT73-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(d)18 tail. The sequence tag for this library is

TGATCAGCT.

TAG-LIB-UI-H-E21

TAG-TISSUE-grade-2-Chondrosarcoma

TAG-SEQ-ATCTAATATG"

BASE COUNT 101 a 213 c 253 g 162 t

ORIGIN

Query Match 43.0%; Score 567.4; DB 14; Length 729;

Best Local Similarity 99.8%; Pred. No. 6.6e-92;

Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

733 GGGCCACGACCGCTGGCGCTGGCGCGCCGCGGAGAGCGCTTCACCGCTTC

Db 729 GGGCCACGACCGCTGGCGCTGGCGCGCCGCGGAGAGCGCTTCACCGCTTC

Qy 813 AGCCACCTATCCCTACTGACGACGAGATGACCTGCACCCACCATCTGCTGAG

Db 669 AGCCACCTATCCCTACTGACGACGAGATGACCTGCACCCACCATCTGCTGAG

Qy 873 ACGGGAGAGAGCCCGCCACCTACGAGGCGCCCTGACAGCTTGGGAGCCCGAGC

609 ACGGGAGAGAGCCCGCCACCTACGAGGCGCCCTGACAGCTTGGGAGCCCGAGC

Qy 933 AGCAGCTGGAACCTGAACCGGAGTGGTGGCGACCCCAAGAACATCTTGAGCA

Db 549 AGCAGCTGGAACCTGAACCGGAGTGGTGGCGACCCCAAGAACATCTTGAGCA

Qy 993 GTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG

Db 489 GTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG

Qy 1053 TCAGCCCGACAGCTGACGAGCGCGCGCATGAGAGGCGCGCGCGCGCGCGCGCG

Db 429 TCAGCCCGACAGCTGACGAGCGCGCGCATGAGAGGCGCGCGCGCGCGCGCGCG

Qy 1113 AGGTATCGGACCTATACCGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT

Db 369 AGGTATCGGACCTATACCGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT

Qy 1173 CCTTCTGAGAGGAGCGCGCTCCACACACATCTGCGCCCTTGAAGAGCGACCA

Db 309 CCTTCTGAGAGGAGCGCGCTCCACACACATCTGCGCCCTTGAAGAGCGACCA

Qy 1233 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 249 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Qy 249 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 249 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Qy 249 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 249 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Qy 249 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 249 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Qy 249 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 249 TCTGAGCAAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Qy 1293 GGGCTGGGCGTGCCTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 189 GGGCTGGGCGTGCCTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

RESULT 15

LOCUS BM677602/c

DEFINITION UI-E-EO1-ald-p-09-0-UI-s1 UI-E-EO1 Homo sapiens cDNA clone

ACCESSION UI-E-EO1-ald-p-09-0-UI 3', mRNA sequence.

VERSION BM677602

KEYWORDS EST.

SOURCE BM677602.1 GI:18987498

ORGANISM human.

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL Bonaldo,M.F., Lennon,G. and Soares,M.B.

MEDLINE Normalization and subtraction: Two approaches to facilitate gene

COMMENT Genome Res. 6 (9), 791-806 (1996)

CONTACT 97044477

CONTACT Contact: Soares, MB

PROGRAM Program for Rat Gene Discovery and Mapping

UNIVERSITY University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

TEL: 319 335 8250

FAX: 319 335 9555

EMAIL: msares@blue.weeg.uiowa.edu

TISSUE Tissue Procurement: Dr. Gregg Hageman

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

Genetics (www.resgen.com).

Seq primer: M13 Forward

POLYA-Yes.

Location/Qualifiers

1. 730

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-E-EO1-ald-p-09-0-UI"

/tissue="Fetal eye"

/dev\_stage="Fetal eye"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a

modified polylinker; Site\_1: EcoR I; Site\_2: Not I;

UI-E-EO1 is a normalized cDNA library containing the

following tissue(s): Fetal eye. The library was

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA,

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pRT73-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(d)18 tail. The sequence tag for this library is

CGGCTATACC. This library was created for the program, Gene

Discovery in the Visual System, supported by National Eye

Institute (NEI).

TAG-LIB-UI-E-EO1

TAG-TISSUE=human fetal eye

BASE COUNT 103 a 210 c 251 g 164 t 2 others

ORIGIN

Query Match

Best Local Similarity 42.8%; Score 564.8; DB 14; Length 730;

Matches 566; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
OY 752 CGGCCACCGACCGCTGGCGGCTGCGGCTTCGCCAGCGGAGGCGCTTCACCGCTTC 811
Db 730 CGGCCACCGACCGCTGGCGGCTGCGGCTTCGCCAGCGGAGGCGCTTCACCGCTTC 671
OY 812 CAGCCACCTATCCGTACCTGACGACGAGATCGACTGCGCACCCACCATCTCGCTGCA 871
Db 670 CAGCCACCTATCCGTACCTGACGACGAGATCGACTGCGCACCCACCATCTCGCTGCA 611
OY 872 GACGGGGAGAGCCGCCACCTACAGGGCCCCGTGACCTTCACGCTTCGGGAGCCCGAG 931
Db 610 GACGGGGAGAGCCGCCACCTACAGGGCCCCGTGACCTTCACGCTTCGNGAGCCCGAG 551
OY 932 CAGCAGCTGAGTGAACCGGAGTGGGTGGCGGCGACCCGCCAAGCAACCATCTTCGAC 991
Db 550 CAGCAGCTGAGTGAACCGGAGTGGGTGGCGGCGACCCGCCAAGCAACCATCTTCGAC 491
OY 992 AGTGACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCAGCAGTAACTCGGGC 1051
Db 490 AGTGACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCAGCAGTAACTCGGGC 431
OY 1052 ATCAGCCGCTACGCTGCTACGGCAGCGGCGGCGCATGAGAGGGCGCGGCCACTACAGC 1111
Db 430 ATCAGCCGCTACGCTGCTACGGCAGCGGCGGCGCATGAGAGGGCGCGGCCACTACAGC 371
OY 1112 GAGGTCACTGGGCACTACCGGGGCTCTCTCCAGCAACGAGAGAGTGGGCGCGCC 1171
Db 370 GAGGTCACTGGGCACTACCGGGGCTCTCTCTCCAGCAACGAGAGAGTGGGCGCGCC 311
OY 1172 TCCTTGCTGAGAGGGGACCGGCTCCACCAACACATCGGCCCCCTAGAGAGCGCAGCC 1231
Db 310 TCCTTGCTGAGAGGGGACCGGCTCCACCAACACATCGGCCCCCTAGAGAGCGCAGCC 251
OY 1232 ATCTGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTAGGTTCCCGAGGGGGC 1291
Db 250 ATCTGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTAGGTTCCCGAGGGGGC 191
OY 1292 CGGGCTGGGGCTGGCTAGGTGAAAGGCA 1321
Db 190 CGGGCTGGGGCTGGCTAGGTGAAAGGCA 161
```

Search completed: March 17, 2003, 13:40:12  
Job time : 2492.96 secs

**THIS PAGE BLANK (USPTO)**